MIRCEA ȘILEȚCHI • LISETTE CURTA

Modelarea matematică a proceselor sociale

Scanned with OKEN Scanner

MODELAREA MATEMATICĂ

A

PROCESELOR SOCIALE

DE

MIRCEA ȘILEȚCHI ȘI LISETTE CURTA

Scanned with OKEN Scanner

CUVÎNT ÎNAINTE

Strălucitul program de acţiune pentru dezvoltarea conştiinţei socialiste, revoluţionare, pentru ridicarea la un înalt nivel a întregii activităţi educative din ţara noastră, marchează în mod indiscutabil un moment de cea mai mare însemnătate în viaţa şi activitatea Partidului Comunist Român şi a întregii noastre societăţi. Ceea ce se degajă cu pregnanţă din acest program este înalta principialitate marxist-leninistă, consecvenţa în aplicarea liniei adoptate de Congresul al X-lea spre a asigura înaintarea fermă a societăţii noastre pe calea făuririi societăţii socialiste multilateral dezvoltate. Fără îndoială, măsurile recente propuse de secretarul general al partidului, tovarăşul Nicolae Ceauşescu, vor determina un adevărat reviriment în activitatea ideologică, cu efecte multiple şi din cele mai favorabile ale aplicării politicii partidului în toate sectoarele de activitate umană.

În contextul acestor măsuri, celor care lucrează în domeniul ştiinţelor sociale le revin sarcini importante pe linia lărgirii continue a ariei de cercetare ştiinţifică în acest domeniu, în vederea înarmării lor cu problemele teoretice şi practice ale ştiinţei conducerii societăţii socialiste multilateral dezvoltate, ştiut fiind faptul că această ştiinţă reflectă fidel dinamismul proceselor sociale din societatea noastră socialistă. De aceea, perfecţionarea metodelor şi modelelor matematice de analiză sociologică a acestor procese deosebit de actuale constituie o preocupare utilă în activitatea extrem de vastă a muncii de cercetare ştiinţifică în ştiinţele sociale, unde problemele conducerii sociale legate de practica construcţiei socialiste din ţara noastră ocupă un loc deosebit de important.

Scanned with OKEN Scanner

În cadrul acestor preocupări s-au întreprins unele cercetări științifice promițătoare, în direcția extinderii și în țara noastră, a aplicațiilor matematicii și a calculatoarelor electronice în domeniul sociologiei. De altfel abordarea teoretică și practică în acest domeniu se bucură de o tradiție recunoscută în țara noastră[1] *și continuată astăzi, atît de matematicieni cît și de sociologi.*

Lucrarea Modelarea matematică a proceselor sociale, *pe care o prezentăm sociologilor din țara noastră își propune să prezinte rezultatele cercetărilor cu caracter metodologic efectuate, în perioada 1967—1970, de Centrul de cercetări sociologice al Academiei de Științe Sociale și Politice, avînd drept obiectiv elaborarea și perfecționarea aplicațiilor posibile ale metodelor și modelelor matematice de analiză a proceselor sociale, în cadrul sociologiei matematice, disciplină de frontieră între sociologie și matematică.*

În vederea realizării acestui obiectiv, am căutat pe cît posibil să aprofundăm unele aspecte teoretice ale aplicațiilor metodelor și modelelor matematice în cercetările sociologice, astfel încît acestea să permită sociologilor, care doresc să le aplice în studiul proceselor sociale, să-și formeze o imagine cît mai completă asupra utilității matematicii într-o activitate de aprofundare a analizei acestor procese deosebit de complexe.

Fără îndoială că în lucrare nu se urmărește să se prezinte exhaustiv toate metodele și modelele de analiză sociologică, ci mai mult se examinează cele solicitate de practica cercetărilor sociologice din țara noastră.

Cu ocazia redactării lucrării s-a căutat ca prezentarea matematică să fie cît mai accesibilă sociologilor cu cunoștințe minime de matematică modernă, urmărindu-se ca expunerea teoretică să fie însoțită de exemplificări numerice și de reprezentări grafice, care să vină în sprijinul unei înțelegeri mai ușoare a lucrării.

Punînd la dispoziția sociologilor această lucrare, reprezentînd un început de drum al autorilor, evident nelipsit de greutăți și lipsuri în abordarea unui domeniu în formare,

[1] Spiru Haret, *Mecanica socială*, București, Edit. științifică, 1962.

sperăm că ea va fi totuși utilă în activitatea lor de cercetare științifică, ajutîndu-i la rezolvarea problemelor multiple pe care le ridică analiza proceselor sociale, caracterizate de un ritm accelerat și de o complexitate crescîndă.

Aducem calde mulțumiri tovarășilor prof. H. H. Stahl, prof. dr. Constantin Ionescu, dr. Ovidiu Bădina, Ioan Matei, Oscar Hoffman, dr. Tiberiu Schatteles, Ion Răvar, Mircea Manolescu, Barbu Berceanu, Honorina Cazacu, Maria Fulea, Maria Cobianu, Doina Gheorghieș, Colectivului de matematicieni de la Centrul de calcul și cibernetică economică și tuturor celor care, prin sugestiile și propunerile lor valoroase, ne-au acordat un ajutor neprețuit la îmbunătățirea acestei lucrări.

Nu putem încheia aceste rînduri înainte de a ne exprima recunoștința și de a aduce un respectuos omagiu profesorilor Miron Constantinescu, Mircea Maliță, Grigore Moisil, Nicolae Teodorescu și H. H. Stahl care, prin preocupările domniilor lor, prin îndrumarea și încurajarea aplicațiilor statistico-matematice în științele sociale, au reprezentat un exemplu demn de urmat și, în același timp, un imbold în realizarea acestei lucrări.

AUTORII

București, 18 septembrie 1971

Scanned with OKEN Scanner

1. PREOCUPĂRI ACTUALE DE MODELARE MATEMATICĂ A PROCESELOR SOCIALE

1.1. INTRODUCERE

Extinderea aplicaţiilor matematicii şi a aplicaţiilor calculatoarelor electronice în ştiinţele sociale, îndeosebi în domeniul sociologiei, ridică o problemă esenţială : în ce măsură aceste aplicaţii pot fi utile pentru abordarea diferitelor etape ale conducerii, planificării şi analizei proceselor sociale? În prezent, la această întrebare, nu se poate răspunde în bloc.

În ultimul deceniu, anumite ramuri importante ale matematicii moderne ca teoria mulţimilor, teoria grafelor, teoria informaţiei, teoria jocurilor etc. au realizat progrese spectaculoase, devenind apte pentru extinderea unor aplicaţii practice în domeniul ştiinţelor sociale. Aceste aplicaţii pot fi rezolvate cu ajutorul calculatoarelor electronice, care sînt capabile de performanţe remarcabile : rapiditate în efectuarea calculelor şi o capacitate mare de înmagazinare a informaţiilor în „memoria" lor.

Programele de cercetare ştiinţifică a proceselor sociale recurg din ce în ce mai mult la aportul matematicii şi al calculatoarelor electronice. Această tendinţă nu se manifestă în mod întîmplător ; ea este impusă de însăşi condiţiile generale ale dezvoltării sociale. Într-adevăr, creşterea continuă a complexităţii vieţii sociale şi a proceselor sociale a determinat cadrele cu puterea de a decide, să recurgă la modele matematice şi la calculatoare electronice, atît pentru fundamentarea deciziilor umane cît şi pentru prevederea consecinţelor pe care acestea le pot avea asupra oamenilor.

Accelerarea proceselor de industrializare, de urbanizare, de dezvoltare a activităţilor agricole etc., cu efecte

Scanned with OKEN Scanner

puternice asupra mobilităţii spaţiale, profesionale şi sociale a forţei de muncă, ne obligă să abordăm aceste probleme cu metode ştiinţifice, apelînd tot mai mult la aplicaţiile matematicii şi ale calculatoarelor electronice, tocmai datorită complexităţii acestor procese.

Cercetarea ştiinţifică modernă abordează procese sociale din ce în ce mai complexe; în asemenea condiţii, ar fi inutil să se încerce izolarea unor relaţii matematice, fără să se ţină seama de întregul ansamblu de relaţii care determină evoluţia acestor procese.

O relaţie dintre două variabile, prinsă într-o expresie matematică, nu poate să reflecte în mod satisfăcător esenţa unui proces social dat. Interacţiunea mai multor variabile este cea care trebuie şi poate să fie exprimată printr-un model matematic rezolvabil la un calculator electronic.

Elaborarea de modele matematice ale proceselor sociale se izbeşte însă de obstacole serioase; pe de o parte, ostilitatea celor care, necunoscînd matematica, nu vor s-o aplice în activitatea lor practică; pe de altă parte, excesul de zel al altora care, avînd înclinaţii spre modernism, mai mult proslăvesc matematica, decît se preocupă de importanţa problemelor abordate şi de traiectoria rezultatelor obţinute [1]. Făcînd chiar abstracţie de dificultăţile de ordin psihologic, cercetătorii ştiinţifici sînt totuşi preocupaţi de utilitatea matematicii şi a calculatoarelor pentru ştiinţele sociale.

Ştiinţele sociale, ca de pildă economia, sociologia, psihologia, istoria etc., se pretează în mod diferit la matematizare. Utilitatea aplicaţiilor matematicii în aceste ştiinţe, depinde, pe de o parte, de natura datelor culese şi de tipul informaţiilor recoltate iar pe de altă parte, de procedeele utilizate pentru a le culege şi de metodele de analiză folosite.

Cu alte cuvinte, interacţiunea dintre tipul de informaţii conţinute în date şi instrumentul matematic utilizat, determină în ultimă instanţă utilitatea aplicaţiilor matematicii în domeniul sociologiei. De pildă, atîta timp cît nu

Scanned with OKEN Scanner

se vor pune de acord informaţiile sociologice (recoltate cu ajutorul chestionarelor de anchetă) cu metodele matematice de analiză a lor, utilitatea aplicaţiilor matematicilor în cercetările sociologice va fi iluzorie.

Perspectivele aplicaţiilor utile ale matematicii şi calculatoarelor electronice în cercetările sociologice sînt strîns corelate atît cu gradul de perfecţiune atins de instrumentul matematic utilizat în aceste cercetări, cît şi cu dezvoltarea sociologiei ca ştiinţă, care să facă posibilă punerea la punct a unor metode proprii de investigaţie sociologică. Mai mult ca oricînd, cercetările cu caracter metodologic efectuate pînă în prezent, scot în evidenţă necesitatea elaborării unor metode şi modele matematice de analiză, predicţie şi de optimizare a proceselor sociale, care să ţină seama de particularităţile acestor procese deosebit de complexe, adică să fie create pentru necesităţile de cercetare sociologică.

1.2. APORTUL STATISTICII MATEMATICE ÎN CERCETAREA PROCESELOR SOCIALE

În studiul nostru ne propunem să examinăm succint noţiunile şi conceptele de bază ale sociologiei matematice, cadrul în care se situează modelarea matematică a proceselor sociale. Acest studiu îl vom consacra descrierii principalelor etape ale unei cercetări de sociologie matematică : formalizarea, dezvoltarea matematică a modelului, confruntarea statistică şi concluziile cercetării [2].

O schemă a principalelor etape ale unei cercetări de sociologia matematică este dată în figura 1.

Analiza sumară a schemei din figura 1 evidenţiază următoarele :

a) formalizarea informaţiilor sociologice obţinute din investigaţiile de teren permite formularea modelului fenomenului sau procesului social analizat ;

Scanned with OKEN Scanner

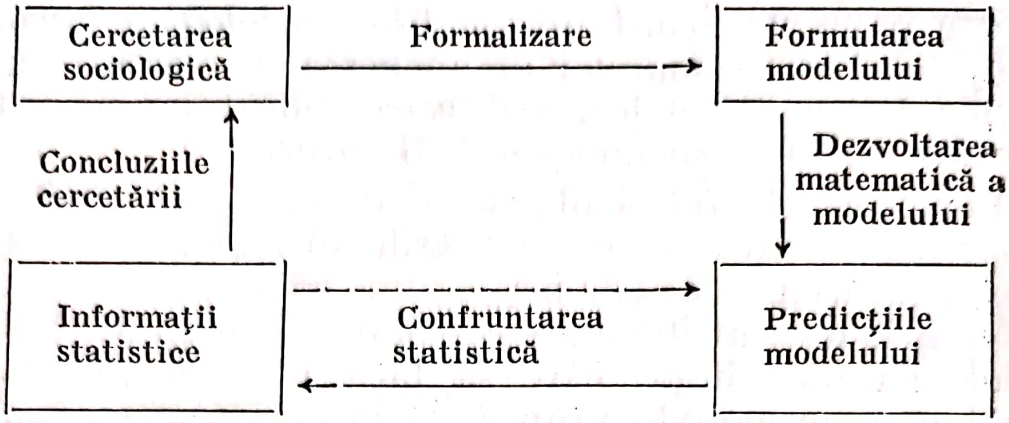


Fig. 1. — Principalele etape ale unei cercetări de sociologie matematică

b) dezvoltarea matematică a modelului formulat din punct de vedere sociologic asigură fundamentarea de predicţii asupra dezvoltării fenomenului sau procesului supus analizei ;

c) confruntarea predicţiilor modelului matematic cu informaţiile statistice obţinute scoate în evidenţă valoarea metodei aplicate în cercetarea sociologică concretă ;

d) concluziile rezultate din analiza informaţiilor statistice culese orientează cercetarea sociologică într-o direcţie nouă, permiţînd aprofundarea analizei unor fapte sociale care prezintă un interes sporit pentru sociologi.

Fiecare etapă a unei cercetări de sociologie matematică ridică probleme deosebit de complexe, a căror rezolvare implică elaborarea unor modele matematice ale proceselor sociale, fără ca, prin aceasta, interesul pentru valoarea lor ştiinţifică să scadă, ci dimpotrivă, să crească.

1.3. FORMALIZAREA INFORMAŢIILOR SOCIOLOGICE

Tehnicile moderne ale modelării proceselor sociale, unele matematizabile iar altele nu, au început să fie aplicate în cercetările sociologice concrete. Aceste tehnici servesc studierii proceselor sociale prin stabilirea unui sistem de caracteristici cărora li se acordă o capacitate de a analiza mai corect decît cea a descrierilor clasice [3]-

Alegerea acestor caracteristici ridică o serie de dificultăţi de ordin metodologic :

a) selecţia caracteristicilor calitative şi cantitative se opreşte de obicei doar la un anumit număr limitat de trăsături ale procesului social analizat, cu o pronunţată notă de arbitrar, întrucît nu avem certitudinea că aceste caracteristici sînt cele mai relevante pentru procesul studiat ;

b) cuantificarea caracteristicilor calitative este deosebit de dificilă ; dificultăţile create de cuantificarea acestor caracteristici le limitează posibilităţile de tratare matematică într-o etapă ulterioară de analiză a procesului social luat în studiu.

Cu ocazia selectării sistemului de caracteristici ale unui proces social este deosebit de util să se efectueze un sondaj sociologic care să permită alegerea obiectivă a caracteristicilor relevante pentru procesul social analizat.

Modelul unui proces social constituie un mijloc ştiinţific de depistare a factorilor determinanţi ai acestui proces deosebit de complex. Pentru a fi stabilită importanţa acestor factori pentru procesul luat în studiu este necesar să introducem, în modelul formulat, factorii cei mai importanţi, în aşa mod încît, ori de cîte ori e posibil, să se asigure eventuala lor cuantificare, care să permită tratarea matematică. Sistemul de caracteristici relevante ale unui proces social reprezintă un ansamblu de ipoteze, deoarece fiecare caracteristică calitativă sau cantitativă luată în parte, reprezintă o ipoteză de lucru [3]. Un astfel de sistem de caracteristici poate constitui o alternativă, el putînd fi foarte bine înlocuit printr-altul, fără a se obţine rezultate sensibil diferite.

Într-o cercetare concretă, sociologul este dator să formuleze un model original al pocesului social luat în studiu, adaptat particularităţilor specifice ale acestuia, în aşa mod încît să constituie un mijloc ştiinţific de investigaţie sociologică. Desigur că, în construirea unui model al unui proces social, pe lîngă unele aspecte specifice ale realităţii studiate, intervin o serie de aspecte de teorie

Scanned with OKEN Scanner

sociologică, rezultate dintr-un proces de raţionament experimental, legat îndeosebi de emiterea ipotezelor de lucru. De fapt, din marea complexitate a acestei realităţi, reţinem unele aspecte, cărora, din punct de vedere teoretic, le acordăm o semnificaţie deosebită.

În realitatea socială, procesele sociale nu sînt izolate, ci fac parte dintr-un sistem complex. Datorită acestui fapt, dificultăţile ridicate de izolarea şi descoperirea proceselor separate, creează greutăţi în activitatea de analiză sociologică. În asemenea condiţii, un model al unui proces social ar fi foarte greu de realizat, chiar dacă l-am simplifica foarte mult. Un asemenea model nu poate fi conceput fără să apelăm la un raţionament deductiv; deci sintetizarea componentelor sistemului funcţional cu scopul de a-i studia transformările în diferite condiţii deşi este de dorit, este greu de realizat.

Modelul unui proces social nu poate fi decît modelul unui sistem ipotetico-deductiv. De pildă, sistemul de caracteristici relevante ale unui asemenea proces este conceput în mod ipotetic ca o structură organic închegată, raţionament logic în lipsa căruia folosirea tehnicii modelării ar deveni inoperantă.

În general, formalizarea reprezintă un procedeu caracterizat tocmai prin explicitarea regulilor de deducţie, potrivit cărora enunţurile pot fi derivate unele din altele. O astfel de formalizare se maturizează prin faptul că în formarea şi derivarea enunţurilor se face abstracţie de semnificaţia lor. Un sistem formalizat este apoi interpretat prin aceea că simbolurilor li se asociază o anumită semnificaţie. De pildă în algebra lui G. Boole, un sistem formalizat, poate primi, între altele, o interpretare matematică şi una sociologică, după cum „1" şi „0" sînt interpretate ca numere sau ca nume pentru „da" şi „nu".

Un cercetător din domeniul ştiinţelor umane trebuie să-şi pună întrebarea: de ce abordăm studiul proceselor sociale pe această cale abstractă, formală, cu ajutorul simbolurilor matematice? Ce legătură au simbolurile

Scanned with OKEN Scanner

matematice folosite pentru descrierea informaţiilor sociologice cu realitatea socială, aşa cum o ştim noi?

Răspunsul cel mai frecvent al oamenilor de ştiinţă este acela că simbolurile matematice servesc doar comodităţii de operare cu ele, reprezentînd un fel de limbaj simplificat, o descriere mai prescurtată, într-un fel mai sintetică, fără de care nu s-ar putea cuprinde şi stăpîni abundenţa informaţiilor sociologice. Esenţial este însă că ele sînt un mijloc deosebit de eficace pentru a pătrunde în realitatea ascunsă dincolo de manifestările, aşa cum le observăm şi percepem noi, ale fenomenelor şi proceselor sociale.

Realitatea socială care îl înconjoară pe un sociolog este pentru el indiscutabilă, după cum nu poate fi pus la îndoială modul de manifestare al oamenilor, cu toate că uneori percepţia îl poate conduce la judecăţi discutabile sau chiar false cu privire la fenomenele sau procesele sociale analizate. Problema implică o soluţionare satisfăcătoare în scopul găsirii modalităţilor de a face afirmaţii obiective, eliberîndu-ne pe cît este posibil de ceea ce este subiectiv, dependent de voinţa noastră.

În domeniul sociologiei, simbolurile matematice nu sînt numere ca în matematică pură, ci ele relevă un fel de realitate socială, care se află ascunsă dincolo de observarea nemijlocită a fenomenelor şi proceselor sociale. Acest fapt poate fi pus în legătură cu întrebarea : cum este posibilă transformarea impresiilor subiective în expresii cu nuanţe obiective?

Simbolurile matematice reprezintă sisteme de caracteristici, dintre cele mai potrivite, destinate descrierii proceselor sociale analizate. Aceste simboluri sînt decisive pentru cunoaşterea obiectivă a proceselor sociale, deoarece cu ajutorul lor se analizează relaţiile cauzale existente, folosindu-se în acest scop un limbaj comun de către toţi cercetătorii care studiază aceste procese deosebit de complexe.

În domeniul sociologiei poate fi stabilită o corespondenţă între simboluri matematice şi percepţii directe. Această corespondenţă se desăvîrşeşte prin asocierea de simboluri matematice unor caracteristici observate. Cerce-

Scanned with OKEN Scanner

tătorii ştiinţifici sînt capabili să pună în evidenţă unele structuri conceptuale camuflate în simbolurile matematice. Aceste simboluri au o particularitate : ele posedă structuri matematice. Deoarece există o asociere între simbolurile matematice şi informaţiile sociologice, structurilor matematice le pot corespunde structuri sociale.

Analiza structurilor sociale implică alegerea atentă a sistemelor de caracteristici şi depistarea relaţiilor cauzale dintre ele. Spre deosebire de caracteristicile calitative sau cantitative, care pot fi observate cu ocazia efectuării unei investigaţii sociologice, depistarea relaţiilor cauzale ridică dificultăţi mari, deoarece aceste relaţii deosebit de complicate sînt neobservabile de către oameni, întrucît ele nu se încadrează în domeniul percepţiilor lor directe şi nu se lasă descrise decît prin concepte abstracte. Datorită acestui fapt, structurile sociale camuflate în simbolurile matematice pot fi înţelese de către sociologi ca realităţi sociale care sînt ascunse în spatele percepţiilor directe. Cu alte cuvinte, structurile sociale pot fi, pe drept cuvînt, imagini ale realităţii sociale, aflate în spatele aparenţelor.

În studiul proceselor sociale nu putem stabili aserţiuni definitive şi categorice. Unele noţiuni, ca valabilitate absolută, exactitate absolută, adevăr definitiv ş.a.m.d. nu-şi găsesc locul în domeniul ştiinţelor sociale. De pildă, procesele sociale nu au o evoluţie deterministă, ci una probabilistă. De cele mai multe ori, dintr-o cunoaştere, totdeauna limitată, a unei evoluţii trecute şi prezente, se pot emite ipoteze verificabile asupra dezvoltării viitoare a proceselor sociale, printr-o exprimare în termeni probabilistici. Din punct de vedere teoretic însă, fiecare aserţiune probabilistă poate fi adevărată, nedeterminată sau falsă.

1.4. DEZVOLTAREA MATEMATICĂ A MODELELOR

Încă cu mai bine de o jumătate de veac în urmă, V.I. Lenin remarca „cucerirea fizicii de către spiritul matematicii" ; acum acest spirit pătrunde şi în domeniul ştiinţelor

umane, adică în economie, sociologie, psihologie, biologie, lingvistică etc. Un indiciu evident, deşi nu întotdeauna bine definit, al prezenţei acestui „spirit" îl poate constitui răspîndirea în ştiinţele umaniste a acelor metode de cercetare care în mod obişnuit sînt considerate exacte. De pildă, limbajul matematic este exact, deoarece termenii lui, ca şi regulile folosirii şi interpretării lui sînt strict determinate din punctul de vedere al matematicii.

Dezvoltarea spectaculoasă a cercetărilor concrete în economie a fost posibilă graţie atît extinderii metodelor şi modelelor matematice, cît şi datorită utilizării maşinilor electronice de calcul. Este evident că dezvoltarea în viitor a cercetărilor concrete de sociologie se va realiza pe aceeaşi cale, adică prin extinderea pe scară largă a modelării matematice şi a tehnicii moderne de calcul.

Metodele cantitative folosite în cercetările sociologice se impun tot mai mult pe plan mondial, ilustrînd legătura tot mai strînsă care se stabileşte între matematică şi sociologie. Aceste metode constituie un mijloc important de cunoaştere ştiinţifică în acest domeniu, prezentînd o deosebită valoare teoretică şi metodologică pentru dezvoltarea sociologiei ca ştiinţă.

Lărgirea necontenită a aplicaţiilor matematicii în domeniul sociologiei determină creşterea eficienţei şi a rolului matematicii în cunoaşterea vieţii sociale, care capătă un conţinut nou în condiţiile extinderii cercetărilor sociologice în etapa actuală.

Interesul crescînd pentru metodele matematice utilizate în cercetările sociologice contemporane are cauze multiple :

— dezvoltarea cercetărilor sociologice în profunzime prin abordarea calitativă şi cantitativă a studiului fenomenelor şi a proceselor sociale ;

— răspîndirea pe scară largă a metodelor şi tehnicilor de cercetare moderne, prin extinderea metodei interviului, sondajului şi analizei informaţiilor obţinute în urma efectuării investigaţiilor în rîndul unor mase mari de oameni ;

Scanned with OKEN Scanner

— perfecţionarea unor teorii matematice care constituie deja capitole importante ale matematicii, cum ar fi : teoria mulţimilor, teoria grafelor, teoria informaţiei, teoria jocurilor, teoria probabilităţilor etc. ;

— extinderea tehnicii electronice de calcul în continuă perfecţionare care permite prelucrarea unui volum mare de informaţii şi aprofundarea analizei matematice a fenomenelor şi proceselor sociale studiate.

În prezent există preocupări în direcţia perfecţionării şi extinderii metodelor matematice moderne în domeniul sociologiei. Desigur că aceste preocupări nu constituie un prinos adus sociologiei, ci un mijloc eficace de abordare a studiului fenomenelor şi proceselor sociale. Traducerea acestor fenomene şi procese în cifre şi simboluri ne permite să abordăm un număr mare de variabile, să confruntăm caracteristicile lor şi să aprofundăm analiza.

Aplicarea matematicii şi tehnicii electronice de calcul este eficientă şi indispensabilă cercetărilor sociologice care apelează frecvent la metodele cantitative de investigare a fenomenelor şi proceselor sociale.

Pe plan mondial, aplicaţiile matematicii şi tehnicii electronice de calcul în domeniul sociologiei sînt orientate spre rezolvarea următoarelor probleme :

— elaborarea unei tipologii de modele matematice ale fenomenelor şi proceselor sociale ;

— organizarea sistemului informaţional capabil să furnizeze datele necesare cercetărilor sociologice concrete ;

— crearea reţelei de centre de calcul care să asigure înregistrarea, prelucrarea şi analiza matematică a informaţiilor sociologice.

Metodele şi modelele matematice care extind sfera analizei acţiunilor şi deciziilor umane vor căpăta o dezvoltare tot mai mare în disciplinele aparţinînd ştiinţelor umane. Această dezvoltare a aplicării matematicii în domeniul sociologiei, care rezultă din nevoile crescînde de obţinere a unor cunoştinţe cît mai exacte despre fenomenele şi procesele sociale, va constitui un cîştig pentru toate ştiinţele umaniste.

Scanned with OKEN Scanner

Într-un sens mai larg, nu există decît o singură matematică care constituie un limbaj universal, un mijloc de comunicare, avînd drept scop studierea relaţiilor şi corespondenţelor [1].

Într-un sens mai restrîns, trebuie făcută distincţie, pe de o parte, între diversele tipuri de matematici şi pe de altă parte, între diferitele aplicaţii ale matematicii în domeniul ştiinţelor sociale. O primă împărţire distinge între matematicile calitative şi cele cantitative.

Matematica ultimelor secole s-a dezvoltat ca o ştiinţă a raporturilor cantitative, a numerelor şi mărimilor [4]. Această matematică cantitativă era pusă în serviciul dezvoltării ştiinţelor naturii şi tehnicii, dar nu şi în slujba ştiinţelor umane. De altfel, matematica clasică nici nu putea asigura un aparat logico-matematic care să satisfacă cerinţele unei abordări ştiinţifice a fenomenelor şi proceselor sociale, deoarece această matematică şi-a dezvoltat baza teoretică în domeniul variabilelor continue, de altfel nepotrivite obiectului de studiu al ştiinţelor umane. De aceea, în ultima jumătate de veac, această înţelegere a matematicienilor a început să fie perimată, categoria cantităţii fiind înlocuită prin categoria structurii; din acest moment, cantitatea, numărul, a încetat de a mai fi regulatorul suprem al matematicilor.

Ceea ce s-a numit multă vreme matematică modernă, adică algebra şi topologia, nu mai este astăzi subordonată în întregime categoriei cantităţii şi conceptelor de continuitate, limită, mulţime infinită etc. Matematica modernă din zilele noastre şi-a dezvoltat baza teoretică îndeosebi în domeniul variabilelor discontinue. Ceea ce se poate remarca este faptul că dezvoltarea teoriilor matematice în domeniul variabilelor discontinue sugerează apropieri surprinzătoare de problemele studiate de ştiinţele umaniste. Astfel, teoria mulţimilor, teoria grafelor şi topologia îşi propun drept scop să stabilească unele relaţii riguroase între clase de indivizi separate unele de altele prin variabile discontinue. Or, discontinuitatea este una din caracteristicile esenţiale ale raportului între ansambluri calita-

Scanned with OKEN Scanner

tive cum sînt cele studiate de ştiinţele despre societate, despre om. Această idee a deschis un cîmp larg de aplicare a matematicilor calitative în studiul fenomenelor şi proceselor sociale, matematici cu perspective excepţionale în domeniul ştiinţelor umaniste. Legătura dintre analiza calitativă şi analiza cantitativă trebuie pusă în evidenţă prin traducerea conceptelor într-un limbaj operaţional. Conceptualizarea reflectă momentul analizei calitative, iar variabilele, indicatorii, reprezintă momentul analizei cantitative. Caracteristicile grupurilor umane pot fi calitative, dihotomice şi cantitative.

Caracteristicile calitative nu implică o ordonare obligatorie a claselor de caracteristici nonmăsurabile şi nonreperabile, ca de exemplu : profesiunea, statutul matrimonial, sexul etc. Stabilirea nomenclaturii de caracteristici calitative implică definirea claselor calitative enumerate printr-o ordonare arbitrară, baza unor raţionamente logice, pe o listă codificată de nomenclatură. Este util să fie standardizate codurile pentru a fi utilizate de către sociologi în mod uniform, deoarece elaborarea unui cod necesită timp şi eforturi mari. De pildă, pot fi elaborate coduri pentru activităţi colective, categorii socioprofesionale, probleme de instrucţie, cauze ale migraţiei, ale fluctuaţiei etc.

Caracteristicile dihotomice pun în evidenţă două categorii care semnalează prezenţa sau absenţa unei caracteristici calitative sau cantitative, ca de exemplu : individul este sau nu este migrant, este sau nu este ataşat de locuinţă etc.

Efectivul unui grup uman studiat poate fi determinat cu ajutorul relaţiei :

$$N = N_1 (X_j) + N_0 (X_j),$$

unde :

N — numărul persoanelor observate ;

$N_1(X_j)$ — numărul persoanelor care posedă caracteristica X_j ;

$N_0(X_j)$ — numărul persoanelor care nu posedă caracteristica X_j.

Probabilităţile de a avea şi de a nu avea caracteristica X_j pot fi determinate cu ajutorul formulelor :

$$P_1(X_j) = \frac{N_1(X_j)}{N} = 1 - P_0(X_j),$$

$$P_0(X_j) = \frac{N_0(X_j)}{N} = 1 - P_1(X_j),$$

unde :

$P_1(X_j)$ — probabilitatea ca grupul uman studiat să posede caracteristica X_j;

$P_0(X_j)$ — probabilitatea ca acelaşi grup uman să nu posede caracteristica X_j.

Probabilităţile de mai sus pot lua valorile :

$$0 \leqslant P_1(X_j) \leqslant 1 \qquad \text{şi} \qquad 0 \leqslant P_0(X_j) \leqslant 1.$$

În clasificarea dihotomică, prezenţa unei caracteristici date poate fi pusă în evidenţă prin 1 iar absenţa ei prin 0, excluzîndu-se posibilitatea de „nedeterminat", din logica „trivalentă".

Caracteristicile cantitative se bazează pe o clasare care se pretează la o măsurare a caracteristicilor în ordine crescîndă sau descrescîndă. Deosebim caracteristici :

— *continue* care sînt variabile susceptibile la o subdiviziune în clase distincte, în scopul măririi preciziei de observare, ca de pildă vîrsta, vechimea în muncă etc. ;

— *discontinue* care variază unitate cu unitate, ca de pildă numărul de copii al unei familii, categoriile de salarii practicate de o întreprindere care variază după categoriile de muncitori etc.

La subdiviziunea în clase de caracteristici se utilizează o scară gradată, avînd clasele ordonate, în mod obligatoriu, în ordine crescătoare sau descrescătoare.

Scanned with OKEN Scanner

Tehnica de cuantificare constituie operația de clasare a informațiilor, în scopul măsurării caracteristicilor calitative, în ordine crescîndă sau descrescîndă.

Prin urmare, tehnicile de cuantificare aplicate în cercetările sociologice operează cu caracteristici cantitative, care pot fi continue sau discontinue.

Tehnicile de cuantificare nu pot fi aplicate decît în analiza informațiilor sociologice susceptibile de a fi măsurate cantitativ, adică traduse în cifre. Aceste cifre trebuie să fie precise, deoarece, în caz contrar, analiza efectuată pe baza lor va fi de la început eronată.

Măsurarea caracteristicilor calitative constituie principalul obstacol care se ridică în fața cuantificării informațiilor sociologice.

Formularea și stabilirea caracteristicilor cantitative care să descrie cu exactitate unele procese sociale particulare constituie încă o problemă extrem de dificilă, care face ca deocamdată, perspectivele cuantificării informațiilor sociologice să rămînă destul de limitate. Aceste dificultăți pot fi evitate, dacă renunțăm la cuantificarea informațiilor sociologice, apelînd la serviciile caracteristicilor calitative, de altfel mai apropiate de problematica cercetărilor sociologice concrete.

În general, culegerea informațiilor se efectuează printr-un test de cîteva întrebări, conceput astfel încît să permită analiza unui ansamblu de caracteristici calitative sistematic reunite. Fiecare întrebare pusă tuturor subiecților populației selectate poate fi concepută astfel încît subiecții să poată răspunde prin „da" sau „nu". Acest mod de formulare al întrebărilor este propriu procedeului răspunsurilor dihotomice. În alte situații, se pot folosi și cuplurile contrare: „adevărat-fals" „favorabil-defavorabil", „de acord — în dezacord" etc.

În majoritatea cazurilor însă, într-un chestionar de anchetă nu putem formula întrebări la care să se răspundă în bloc prin „da" sau „nu", dacă dorim să obținem informații sociologice cît mai nuanțate. Mai mult chiar, putem

Scanned with OKEN Scanner

să formulăm întrebări deschise oricărui răspuns, aşa cum el este conceput de către subiectul intervievat.

Atunci cînd apelăm la serviciile interviului formal, de obicei folosim un chestionar de anchetă, care poate fi privit ca un ansamblu de caracteristici calitative sau cantitative, formulate precis şi sistematic, reunite în scopul culegerii informaţiilor pentru descrierea fenomenelor sociale luate în studiu. Prezenţa sau absenţa unor astfel de caracteristici poate fi scoasă în evidenţă prin cuplurile: „da — nu" sau „1—0". De data aceasta, nu este obligatoriu pentru subiectul intervievat să răspundă prin „da" sau „nu" la o întrebare concretă, investigatorul social putînd să evidenţieze prezenţa sau absenţa unei caracteristici date la persoana respectivă.

Prin urmare, caracteristicile dihotomice sînt foarte utile investigaţiilor sociologice, unde limitele cuantificării sînt evidente. Avantajul acestor caracteristici este amplificat de faptul că ele pot fi doar calitative, cantitative sau chiar calitativ — cantitative.

Accentul asupra tipurilor de caracteristici aplicabile în cercetările sociologice este apreciat în legătură cu tehnicile matematice de „modelare" sociologică. Aceste tehnici constau tocmai în a studia un proces social prin stabilirea unui ansamblu de caracteristici, denumite „variabile" în matematică şi „indicatori" în statistică.

Tipurile de caracteristici cu care operează sociologii pot fi abordate atît de matematica calitativă cît şi de matematica cantitativă. Mai mult chiar, diferitele tipuri de matematici pot fi diferite ţinînd seama de natura variabilelor folosite. De pildă, matematica calitativă poate fi definită ca matematica care operează îndeosebi cu variabilele nenumerice şi ordinale, iar matematica cantitativă poate să opereze cu variabile ordinale şi, mai ales, cardinale.

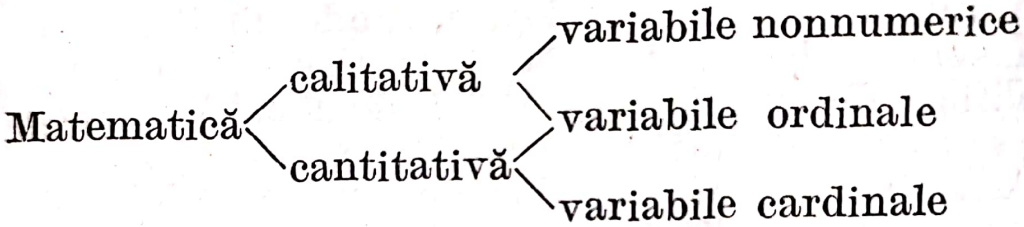
Cu alte cuvinte, matematica calitativă operează cu caracteristici calitative şi dihotomice, iar matematica cantitativă foloseşte caracteristici cantitative şi dihotomice; de aici rezultă valoarea excepţională a caracteristicilor dihotomice pentru cercetările sociologice, deoarece

Scanned with OKEN Scanner

aceste caracteristici elimină deficienţele generate de tehnicile de cuantificare şi favorizează extinderea aplicaţiilor matematicii calitative şi cantitative în domeniul sociologiei.

Matematica modernă nu este cantitativă prin excelenţă, ci ea este mai ales calitativă. Cantitatea nu este pentru matematică decît un element minor. Elementul de bază prin care se construieşte realitatea socială este prin excelenţă calitativ. În timp ce opoziţia dintre calitate şi cantitate este deosebit de pregnantă în domeniul ştiinţelor umane, ea îşi pierde însemnătatea în domeniul matematicii, unde analiza cantitativă nu este esenţială.

Variabilele utilizate în aplicaţiile matematicii în domeniul ştiinţelor umane pot fi sintetizate astfel :



Examinînd succint această schemă putem observa că atît matematica calitativă nonnumerică şi ordinală cît şi matematica cantitativă numerică şi ordinală pot fi utilizate în cercetările sociologice. Cu alte cuvinte, atît caracteristicile cuantificabile cît şi cele necuantificabile, pot servi extinderii aplicaţiilor matematicii în acest domeniu.

De un interes excepţional se bucură aplicaţiile matematicii calitative şi ale matematicii cantitative care implică utilizarea variabilelor cardinale în cercetările sociologice, fără a apela la tehnicile de cuantificare în studiul proceselor sociale.

Matematica aplicată în ştiinţele umane este calitativă prin excelenţă. Această matematică, pe care nici sociologii, nici matematicienii nu ştiu exact unde s-o caute, cu toate că poate fi găsită la tot pasul în studiul realităţii sociale, oferă observaţiilor lor o formă riguroasă, mai precisă [1]. Atît sociologii, cît şi matematicienii vor deopotrivă să scape

Scanned with OKEN Scanner

de dificultăţile generate de tehnicile de cuantificare a informaţiilor recoltate.

De altfel, măsurarea caracteristicilor observate nu constituie un obiectiv indispensabil al sociologilor, dacă avem în vedere faptul că de multe ori cheltuielile mari pentru culegerea unor date cuantificabile, care nu pot indica diferenţe calitative importante, nu sînt justificate. De exemplu, apariţia unui copil într-un menaj implică mai

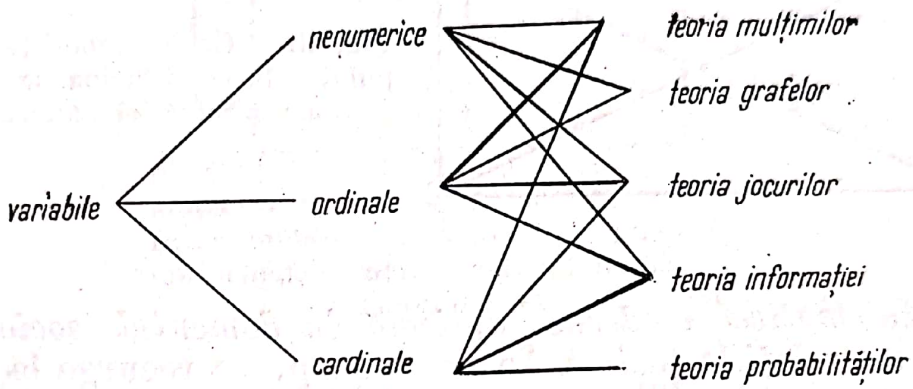


Fig. 2. — Teorii matematice aplicate în domeniul ştiinţelor umane

multe schimbări decît o creştere cu $10^0/_{00}$ a numărului locuitorilor unui oraş. Aici caracteristica calitativă este mai importantă decît caracteristica cantitativă.

Matematica calitativă relevă o noţiune de ordine. Datorită acestui fapt, importantă pentru ştiinţele umane este algebra modernă, care permite compararea unor obiecte diferite dîndu-le o expresie comună.

Teoriile matematice elaborate în ultimul deceniu au recurs la diverse tipuri de matematici, şi anume (fig. 2):

a) teoria grafelor şi teoria jocurilor recurg la variabile numerice şi ordinale;

b) calculul probabilităţilor şi teoria jocurilor apelează la variabilele cardinale;

c) teoria mulţimilor apelează la variabile nenumerice, ordinale şi cardinale.

Aplicaţiile teoriei grafelor au un interes deosebit pentru

cercetările sociologice concrete. Un graf se prezintă ca o configurație a nodurilor (sau a punctelor) și a arcelor (liniilor orientate), cărora li se acordă o semnificație concretă. De pildă, schema unei structuri sociale poate fi reprezentată prin noduri (grupurile sociale) și prin arce (relațiile cauzale dintre grupurile analizate).

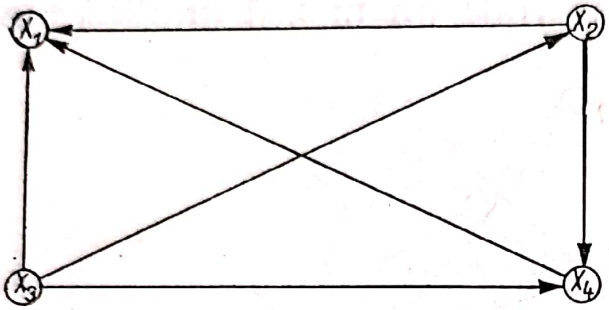


Fig. 3. — Graful mobilității populației între principalele ramuri ale producției materiale

Matematica modernă aplicată în domeniul sociologiei studiază relațiile sociale. În acest scop, ea recurge la algebră, care comportă o structură ușor de reținut și un aparat conceptual deosebit de bogat. De exemplu, între salariații care activează în industrie (X_1), construcții (X_2), agricultură (X_3) și alte ramuri (X_4), poate exista un transfer important de populație după graful dat în figura 3[1].

Acestui graf i se poate asocia o matrice (un tabel) cu dublă intrare. Se scrie 1 dacă există un transfer de populație, și 0 dacă nu există, ca în matricea 1.

Teoria grafelor nu poate să soluționeze toate problemele pe care le ridică modelarea matematică a proceselor sociale. Ea poate cel mult să ne procure un fel de scheme cuantificabile a unor structuri sociale, a relațiilor interpersonale, a circuitelor de tranziție sau de informație.

Teoriile matematice enumerate mai sus pot fi utilizate în studiul rețelelor informaționale, în cercetările

---

[1] Graful prezintă numai rezultatele mișcărilor de salariați, în realitate având loc mișcări în ambele sensuri.

De la \ La	x_1	x_2	x_3	x_4
x_1	0	0	0	0
x_2	1	0	0	1
x_3	1	1	0	1
x_4	1	0	0	0

Matricea 1. — Matricea asociată grafului mobilității populației, între principalele ramuri ale producției materiale

operaționale și în statistică, și anume (fig. 4) : în studiul rețelelor informaționale teoria mulțimilor, teoria grafelor și teoria informației ; în cercetările operaționale teoria mulțimilor, teoria grafelor, teoria jocurilor și teoria probabilităților ; în statistică teoria mulțimilor, teoria jocurilor și teoria probabilităților.

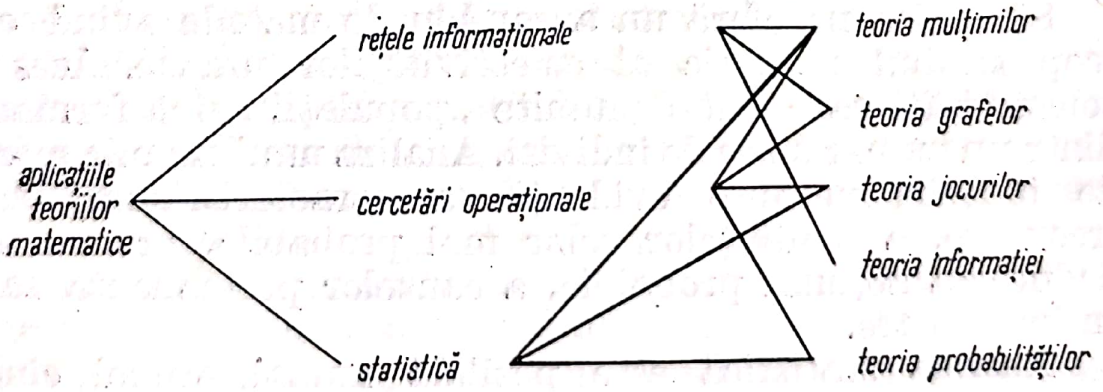


Fig. 4. — Aplicațiile teoriilor matematice

Aplicațiile teoriilor matematice în domeniul științelor umaniste se pot concretiza în următoarele :

a) studiul reţelelor informaţionale care implică măsurarea cantităţii de informaţie furnizată de o anchetă pe bază de chestionar, analiza circulaţiei informaţiilor şi a reţelelor de comunicaţie aferente unor structuri sociale deosebit de complexe etc. ;

b) cercetările operaţionale care permit obţinerea prin metodele logicii matematice a unei scheme de organizare optimă a activităţilor umane din diverse sectoare de activitate ;

c) cercetarea informaţiilor statistice care permite analiza evoluţiei unor fenomene sau procese sociale prin studiul indicilor statistici.

Eficacitatea aplicaţiilor matematicii în domeniul ştiinţelor sociale poate fi sporită prin transformarea datelor calitative în date cantitative, ceea ce măreşte gradul de exactitate al rezultatelor şi permite efectuarea analizei matematice a faptelor sociale cu ajutorul calculatoarelor electronice.

1.5. CONFRUNTAREA PREDICŢIILOR MODELULUI MATEMATIC CU INFORMAŢIILE STATISTICE

Fenomenele şi procesele sociale pot fi studiate în mod ştiinţific prin mijlocirea informaţiilor statistice.

Statistica ne oferă un ansamblu de metode avînd ca scop studiul numeric al colectivităţilor umane. Aceste colectivităţi care sînt denumite „populaţii" sînt formate dintr-un număr mare de indivizi. Analiza unui număr mare de indivizi permite evidenţierea caracteristicilor mai frecvente, a tendinţelor celor mai probabile, precum şi deducerea acţiunii probabile, a cauzelor permanente sau întîmplătoare.

Analiza statistică este posibilă numai atunci cînd caracteristicile fenomenului sau procesului luat în studiu sînt cuantificabile, adică atunci cînd pot fi măsurate într-un fel sau altul. Aceste caracteristici pot fi ordonate astfel încît să formeze serii dinamice, avînd valori crescînde sau

Scanned with OKEN Scanner

descrescînde. Ansamblul de caracteristici ordonate va trebui supus unei prelucrări ulterioare care să demonstreze că seria de valori are o expresie matematică. În faza finală a cercetării, expresia matematică obținută va fi interpretată în termeni sociologici. Informațiile statistice permit trecerea de la nivelele de abstractizare cele mai joase la cele mai înalte prin următoarele etape :

a) analiza statistică a fenomenului sau procesului social luat în studiu, din care se reţin toate aspectele esenţiale ale colectivităţii umane analizate, ale fiecărui individ şi ale tuturor indivizilor ansamblului ;

b) relevarea conexiunilor structurale, funcţionale şi în cele mai multe cazuri cauzale şi semnificative ale fenomenului sau procesului analizat ;

c) concentrarea tuturor aspectelor individuale, multiple şi aparent disparate, într-o expresie matematică, adică într-un *model matematic al fenomenului sau procesului social luat în studiu* ;

d) elaborarea previziunilor sociale cu ajutorul unor procedee de modelare matematică a fenomenelor sau proceselor supuse analizei.

Ca atare, analiza statistică bazată pe expresii matematice constituie primul pas de înţelegere, cunoaştere şi de previziune ştiinţifică a dezvoltării sociale ; tehnicile modelării matematice au însă o capacitate de analiză mai profundă decît cea a metodelor statistice clasice.

Modelarea matematică are un scop ceva mai larg decît o simplă analiză statistică şi anume o modalitate de depistare a factorilor determinanţi ai fenomenului sau procesului social analizat. Mai mult chiar, modelele matematice constituie un mijloc de previziune socială, adică de predicţie a tendinţelor de evoluţie în perspectiva desfăşurării fenomenului sau procesului luat în studiu.

În fine, modelarea matematică ca tehnică de investigaţie ştiinţifică va trebui să contribuie la *perfecţionarea metodologiei de planificare socială* prin extinderea cercetării operaţionale în domeniul sociologiei, care să permită luarea unor decizii optime de dezvoltare a fenomenelor

Scanned with OKEN Scanner

şi proceselor sociale, avînd în vedere obiectivele de ansamblu urmărite de către organele superioare de conducere şi de planificare. Astfel aceste organe ar putea asigura o „dirijare" pe bază ştiinţifică a fenomenelor şi proceselor sociale, îndeosebi prin extinderea utilizării în acest scop a tehnicii de calcul electronic.

Predicţiile de dezvoltare a proceselor sociale elaborate cu ajutorul tehnicilor de modelare matematică trebuie să fie confruntate cu informaţiile statistice observate, în vederea constatării eficacităţii acestor tehnici. Concordanţa predicţiilor modelului matematic cu informaţiile statistice obţinute ne întăreşte încrederea că acest model este apt a fi utilizat în activitatea de planificare socială.

În timpul desfăşurării unui proces social dat se ivesc situaţii care nu mai corespund planificării iniţiale, fapt care implică actualizarea modelului la noile condiţii. Actualizarea modelului matematic constituie operaţia de constatare a factorilor importanţi care în faza iniţială au fost omişi sau neglijaţi, a modului de realizare a sarcinilor de plan la o anumită dată şi de stabilire a noilor obiective de realizat.

Această actualizare a modelului matematic trebuie să se facă periodic, la intervale de timp care să permită o urmărire corespunzătoare a modului de realizare a sarcinilor de plan (de regulă anual, la 5 şi la 10 ani) şi îndeosebi să asigure timpul necesar pentru luarea unor măsuri concrete.

Adaptarea modelului matematic la noile condiţii este posibilă graţie flexibilităţii lui la modificările introduse şi vitezei de calcul atinse de calculatoarele electronice.

1.6. CONCLUZIILE CERCETĂRILOR DE SOCIOLOGIE MATEMATICĂ

Concluziile cercetărilor de sociologie matematică se concretizează în analiza relaţiilor dintre variabile.

Relaţia dintre două variabile X şi Y poate fi :

a) relaţie de dependenţă (Y depinde de X sau invers);

b) relaţie de interdependenţă (Y depinde de X şi X depinde de Y);

c) relaţie de tranziţie (X se transformă — parţial sau total — în Y sau invers);

d) corelaţie statistică (pentru X există întotdeauna Y şi invers);

e) relaţie stocastică (dacă X atunci foarte probabil Y sau invers);

f) relaţie de transmitere a informaţiilor (Y recepţionează informaţii de la X sau invers);

g) relaţie cronologică (dacă X atunci mai tîrziu Y sau invers);

h) covarianţă (dacă X atunci în acelaşi timp Y şi invers).

În domeniul vieţii sociale cele mai răspîndite relaţii sînt relaţiile de interdependenţă, de dependenţă şi corelaţiile statistice [5].

Relaţia de dependenţă este relaţia în care o modificare a unei variabile duce la o modificare a celei de-a doua variabile.

Relaţia de interdependenţă este acea relaţie în care o modificare a unei variabile duce la o modificare a celei de-a doua variabile, care la rîndul ei permite o nouă modificare a primei variabile şi acest proces continuă. Putem obţine modificări mari într-o relaţie de interdependenţă printr-o acţiune redusă asupra uneia dintre variabile.

Corelaţia statistică este relaţia reciprocă între două variabile, din care una apelează logic la alta, semnalînd o asociere dintre ele.

Relaţiile de dependenţă dintre variabilele explicite ale unei structuri sociale date pot fi puse în evidenţă cu ajutorul analizei corelaţionale şi al analizei cauzale. De aceea, analiza corelaţiilor statistice şi a inferenţelor cauzale dintre variabilele unei astfel de structuri prezintă o

Scanned with OKEN Scanner

valoare excepţională pentru cercetarea cantitativă a fenomenelor şi proceselor sociale [2].

În general, stabilirea intensităţii relaţiei dintre două variabile nu ridică probleme complicate, analiza inferenţelor cauzale însă generează dificultăţi, deoarece raportul dintre două variabile se poate modifica fără să existe o relaţie de la cauză la efect. Pentru a stabili relaţii cauzale trebuie să apelăm la serviciile analizei multivariate aplicată mai multor variabile interdependente.

Între o măsură a corelaţiei statistice şi o măsură a inferenţei cauzale există deosebiri evidente. De obicei, o măsură a corelaţiei pune în evidenţă o asociere, un paralelism, dintre două variabile, pe cînd o relaţie cauzală implică eliminarea efectelor „perturbante" ale altor variabile care pot genera o inferenţă cauzală aparentă, falsă [6].

Ordonarea cronologică a variabilelor, îndeosebi a caracteristicilor de opinie ale oamenilor, ridică dificultăţi extrem de mari. Numai dacă se efectuează investigaţii sociologice pe o perioadă lungă există posibilităţi să se stabilească o succesiune cronologică a variabilelor luate în studiu. Evident, tehnicile de eşantionare „panel"[3] pot fi deosebit de utile în acest caz, fără ca să rezolve definitiv această problemă deosebit de complicată.

Influenţa perturbantă a altor variabile poate fi pusă în evidenţă prin cercetarea raportului dintre variabilele în care se poate măsura efectul „manipulării" variabilelor independente asupra variabilelor dependente [7]. Dacă numărul variabilelor perturbante potenţiale creşte foarte mult, analiza relaţiilor cauzale se complică, generînd dificultăţi care pot fi evitate de către cercetători doar prin experimente controlate. Numai aceste experimente permit verificarea relaţiilor cauzale. Cu toate acestea, experimentarea controlată pe populaţii umane fiind greu de realizat,

---

[2] Prin inferenţă înţelegem operaţia logică de trecere de la un concept la altul, în care ultimul este dedus din primul.

[3] Panelul este utilizat de sociologia contemporană în investigarea fenomenului de schimbare în timp, din viaţa socială.

Scanned with OKEN Scanner

posibilităţile de formulare a concluziilor cauzale sînt extrem de limitate.

O sociologie experimentală poate să apeleze la „probe" prin experimente, fără să efectueze o experimentare directă, ca de pildă în fizică, chimie etc. Experimentului indirect i se poate da un sens de probă a unei ipoteze care caută să unească doi factori într-o relaţie cauzală, printr-un studiu al situaţiilor contrastante, ai căror factori au fost controlaţi în întregime, cu excepţia unuia care ne interesează şi care este cauza ipotetică, ori efectul ipotetic [8].

În general, niciodată nu putem fi siguri că o corelaţie statistică dintre două variabile explicite nu se datorează uneia sau mai multor variabile necontrolate, poate chiar nebănuite. Din acest fapt, inferenţele cauzale nu pot fi formulate decît aproximativ, cu multă precauţie. Concluziile trase din analiza *ex post facto* nu pot fi decît speculative, ele putînd da indicaţii mai puţin sigure despre efectele perturbante ale altor variabile decît experimentele controlate avînd drept scop verificarea ipotezelor emise asupra relaţiilor cauzale.

Concluziile cercetărilor de sociologie matematică vor fi obţinute prin verificarea ipotezelor emise asupra sistemelor de variabile şi asupra relaţiilor cauzale dintre ele, cu ajutorul tehnicilor de modelare matematică a fenomenelor şi proceselor sociale.

Tehnicile de modelare matematică nu prezintă o noutate pentru ştiinţă, însă aplicarea lor în domeniul sociologiei este de dată recentă. Fiind un instrument important de cunoaştere a proceselor sociale, aplicarea acestor tehnici în acest domeniu a început să depăşească cadrul cercetării teoretice, devenind, treptat, o preocupare de ordin practic.

Reprezentările matematice au o importanţă teoretică şi practică, deoarece, prin intermediul unor elemente constructive ale unor modele, pot fi create pentru sociologi unele posibilităţi de pătrundere în domenii de cunoaştere greu accesibile, aflate dincolo de limitele observaţiei nemijlocite ale cercetătorului.

Scanned with OKEN Scanner

Aplicarea modelării matematice în domeniul sociologiei implică realizarea unei permanente relaţii între teorie, model şi realitate. Într-o astfel de relaţie, corespondenţa între teorie şi model se prezintă sub aspect *logic*, iar cea între model şi realitate, sub aspect *gnoseologic*. De aceea, modelarea matematică a fenomenelor şi proceselor sociale trebuie dezvoltată şi perfecţionată astfel încît să asigure [9]:

— corespondenţa între teorie şi model;

— corespondenţa între model şi realitate.

Asigurarea acestor două condiţii este deosebit de importantă, deoarece în felul acesta între teorie şi practică se va putea realiza o corespondenţă perfectă, care poate asigura succesul modelării matematice în domeniul sociologiei.

Modelele matematice ale proceselor sociale conduc la un ansamblu de predicţii asupra structurii informaţiilor sociologice; aceste predicţii pot fi confirmate sau infirmate de datele statistice. De pildă, un model matematic al procesului de migraţie a populaţiei de la sat la oraş, poate constitui un ansamblu ordonat de ipoteze verificabile prin confruntare cu rezultatele experimentale. În asemenea condiţii, datele statistice vor fi puse de acord cu predicţiile, doar în condiţiile în care ipotezele de bază ale modelului vor fi confirmate prin experienţă. Numai în aceste condiţii o parte din ipotezele emise pot servi la formularea predicţiilor.

Cercetările de sociologie matematică oferă o clasă de modele mai generale, care permit formularea de predicţii confirmate prin date şi efectuarea unor analize a diverselor particularităţi a acestor modele, în scopul diferenţierii de noi predicţii.

Tehnicile moderne ale modelării matematice au drept scop, în afara analizei descriptive, şi îmbogăţirea explicării fenomenelor şi proceselor sociale, a factorilor şi a relaţiilor cauzale determinante.

Scanned with OKEN Scanner

1.7. APLICAȚIILE TRADIȚIONALE ALE MATEMATICII ÎN ȘTIINȚELE SOCIALE

Economia și demografia sînt primele dintre științele care au utilizat metodele matematice, ele fiind urmate începînd cu ultimii ani ai secolului al XIX-lea de psihologie, iar după cel de-al doilea război mondial, de sociologie.

Aplicațiile și utilizarea matematicii în științele sociale au cunoscut o istorie mai însuflețită decît în științele naturii.

În perioada premodernă, în aplicațiile matematicii în științele sociale, se observă o despărțire între teorie și inducție; cu riscul unei exprimări schematice, se poate spune că teoria rămîne foarte adesea speculativă, în timp ce inducția este oarbă.

Se pot distinge trei căi sau trei tipuri de aplicare a matematicii în științele sociale — proprii acestei perioade premoderne [10]:

— o tradiție teoretică care este cînd normativă, ca de exemplu în lucrările de la sfîrșitul secolului al XVIII-lea în Franța ale lui Buffon, Laplace, Condorcet, care creează cadrul a ceea ce numim astăzi cercetări operaționale; cînd speculativă, ca de exemplu în lucrările lui Cournot în măsura în care aceste metode matematice teoretice nu sînt legate, în general, de verificarea prin comparație cu experiența sau cu observația. Funcția acestor metode constă în a analiza consecințele axiomelor cu ajutorul cărora să se exprime într-o formă idealizată unele fapte sociale reale. O moștenire a acestei tradiții o constituie teoria jocurilor, teoria imitației a lui Raschewsky etc.;

— cea de-a doua tradiție se referă la lucrările considerate astăzi ca aparținînd domeniului statisticii matematice. Principalii pionieri ai acestor tradiții sînt Laplace, Verhulst, Poisson, Lexis ș.a. Este o tradiție de o mare importanță pentru că furnizează exemple de teorie matematică cu valoare inductivă;

Scanned with OKEN Scanner

— în sfîrşit, cea de-a treia tradiţie în acest ansamblu de încercări de a introduce matematica în analiza fenomenelor sociale este cea desemnată de anglo-saxoni prin expresia *curve-fitting*[4]. Obiectivul cercetat de acest tip de aplicaţii este îndeosebi făurirea unui instrument de predicţie : dacă se observă, de exemplu, că un proces cercetat pe o perioadă destul de lungă este ajustat de un polinom de gradul trei, se va utiliza această funcţie pentru a prevedea evoluţia ulterioară a procesului social.

În general, se poate afirma că în sociologia şi psihologia socială metodele matematice îşi găsesc o largă aplicare în problemele dinamicii de grup, la analiza grupurilor umane, în studiul fenomenelor şi proceselor sociale. De pildă, procesul de migraţie a populaţiei de la sat la oraş poate fi analizat matematic.

Unul din domeniile care se pretează cel mai bine utilizării metodelor matematice sînt cercetările asupra grupurilor mici şi cele privind prelucrarea sondajelor sociologice.

Dacă, timp de mulţi ani, matematica utilizată în ştiinţele sociale a constat fie în aplicaţiile metodelor algebrice elementare, fie în analiza matematică clasică, fie în aplicaţiile elementare ale calculului probabilităţilor etc., astăzi o serie de metode moderne ale matematicii pătrund progresiv în cercetările sociale. Acestea la rîndul lor, ajunse la un anumit grad de complexitate, impun crearea de noi instrumente matematice.

Dintre ramurile noi ale matematicii — cu aplicaţii în ştiinţele sociale — se numără algebra superioară, îndeosebi algebra liniară. De pildă, în economie au fost introduse matricele lui L e o n t i e f f iar în psihologie a fost aplicată analiza factorială a matricelor de corelaţie. Ecuaţia de bază în analiza factorială şi în numeroase metode de analiză dimensională utilizate în psihologie este ecuaţia matricială :

$$R = A \cdot A',$$

[4] Prin *curve-fitting* se înţelege curba de ajustare a evoluţiei unui fenomen sau proces care reflectă *trendul* (tendinţa, orientarea, direcţia acestei evoluţii).

Scanned with OKEN Scanner

unde :

R — matricea de corelaţie (de exemplu între probele psihometrice);

A — matricea exprimînd corelaţiile între probele psihometrice şi factorii necunoscuţi sau latenţi (cantităţi care poartă denumirea de „situaţii");

A' — transpusa matricei A.

De asemenea, în psihologia socială, teoria măsurii, metoda de analiză a structurilor latente pun matematicianului probleme algebrice complexe.

În sociologie un mare număr de fenomene dau loc în mod natural la reprezentări sub formă matricială şi ca urmare, aplicaţiilor algebrei matriciale : tabele de mobilitate, tabele de schimbări în caracteristicile observate, cum sînt cele culese prin sondajele panel. Aceste aplicaţii îşi dovedesc utilitatea în studiul tendinţelor comportamentului uman, în analiza atitudinilor.

Aplicaţiile teoriei proceselor stocastice în ştiinţele sociale sînt de dată recentă. Chiar domeniul economiei, care se pretează cel mai bine la aplicaţiile matematicii, nu a apelat decît într-o măsură foarte redusă la această teorie. În domeniul sociologiei sînt utilizate lanţurile M a r k o v , care sînt caracterizate prin faptul că probabilitatea p_{ij} de a trece din starea i în starea j, oricare ar fi i şi j, este dependentă de timp.

În afara lanţurilor Markov, aplicaţii ale teoriei proceselor stocastice se întîlnesc în analiza distribuţiilor descendente, în analiza fenomenelor şi acţiunilor spontane, ca de pildă procesele P o i s s o n .

Datorată matematicianului german K o e n i g şi dezvoltată în continuare, prin lucrările lui C. B e r g e , teoria grafelor constituie un mijloc util de rezolvare a anumitor probleme aparţinînd cercetărilor operaţionale.

În psihologia socială întîlnim numeroase aplicaţii în analiza sociogramelor şi în general a structurilor sociale. Rezultate interesante se obţin prin aplicarea metodelor algebrice matricelor asociate sociogramelor etc.

Scanned with OKEN Scanner

Metodele simulării, prin supleţea lor, pot aduce contribuţii mari în ştiinţe ca economia şi sociologia, îndeosebi unde experimentarea este imposibilă, complexitatea ipotezelor făcînd traducerea matematică foarte dificilă.

Nu este exagerat dacă se afirmă că ştiinţele umane au exercitat o acţiune modestă dar reală asupra dezvoltării anumitor ramuri ale matematicii. Să luăm exemplul teoriei grafelor care născută din lucrările lui K o e n i g (publicate în 1936) sînt trecute relativ neremarcate pînă la începutul deceniului 1951 — 1960, dată la care îşi găseşte primele sale aplicaţii în lucrările lui R o s s şi H a r a r y, privind aplicaţiile sociometrice ale teoriei grafelor.

Se pot da şi alte exemple de teorii care născute mai întîi în contact cu problemele puse de ştiinţele umanistice au devenit adevărate ramuri ale matematicii; acest fenomen este tipic pentru teoria jocurilor, care ridică probleme matematice numeroase. De pildă, problemele teoretice cele mai importante care se pun astăzi în teoria jocurilor sînt cele legate de definirea acţiunilor raţionale în situaţii complexe proprii jocurilor non-competitive cu n persoane (de sumă nenulă), adică jocuri care au cele mai mari şanse de a exprima în mod veridic realitatea socială.

1.8. DIFERENŢIEREA FUNCŢIILOR MATEMATICII

În ultimii ani, asistăm la o diferenţiere accentuată a funcţiilor matematicii, ca rezultat al extinderii aplicării ei în ştiinţele economice şi sociale.

Dezvoltarea tehnicilor de observaţie, în acelaşi timp cu creşterea numărului unităţilor observate, a dat naştere într-o serie de discipline, ca sociologia şi psihologia, unei multitudini de metode şi tehnici a căror funcţie principală constă în a permite cercetătorului să se orienteze într-un ansamblu de informaţii a căror înţelegere depăşeşte posibilităţile sale intuitive. Aşa sînt : analiza factorială asociată numelor lui S p e a r m a n, T h u r s t o n e, H o t e l l i n g,

analiza ierarhică datorată lui Guttman, tehnici de clasificare sau de construcţie de tipologii, avînd drept obiectiv să se substituie intervenţiei intuitive care se foloseşte în viaţa obişnuită şi în anumite probleme ştiinţifice. Aceste tehnici vizează să reproducă logica acestor căi intuitive, îndepărtînd riscurile unei rigori ştiinţifice întîlnite în intuiţie.

O altă funcţie mai recentă a metodologiei matematice este aplicarea sa din ce în ce mai intensă la problemele de experimentare. Aici este vorba de ştiinţele umaniste capabile să organizeze experienţe pe obiectele lor, psihologia şi psihologia socială. Există de exemplu, în psihologia socială o mulţime impresionantă de experienţe şi de lucrări asupra formării şi schimbării opiniilor, atitudinilor sau sentimentelor individuale în funcţie de schimbările care intervin în grupul sau în mediul de care aparţine subiectul.

Începînd cu lucrările lui Heider şi Festinger îndeosebi, se insistă asupra faptului că individul constrîns să-şi modifice unele sentimente sau unele atitudini ale sale, atunci cînd el se găseşte într-o situaţie în care unanimitatea este împotriva lui, recurge la o reorganizare generală a atitudinilor şi a opiniilor sale corelative. Pînă nu de mult aceste teorii (*theory of cognitive organisation* a lui Heider şi *theory of cognitive disonance* a lui Festinger) au favorizat efectuarea unor experienţe de tip clasic. Cartwright şi Harary au formulat aceste teorii cu ajutorul grafelor, reprezentînd noţiunea de echilibru prin noţiunea formală de *graf echilibrat*. Această formulare, deşi foarte elementară, are fără îndoială un dublu aspect : de a preciza conceptele conţinute în teoriile arătate mai sus şi de a formula un cadru experimental precis.

Formularea matematică a fenomenelor şi problemelor sociale a condus la intensificarea experimentării în domeniul sociologiei.

În afară de aceste funcţii noi ale matematicii se mai pot cita şi altele, ca de pildă : cea de *cvasiexperimentare* care se aplică îndeosebi acolo unde nu sînt posibile experienţele, cum sînt domeniile aparţinînd economiei şi sociologiei. În aceste cazuri, este posibil să se substituie realită-

ţii sociale o versiune simplificată, un model matematic cu care se va putea studia comportamentul, variind condiţiile structurale ce i se aplică. Această situaţie este tipică economiei matematice unde nu se poate în general lucra experimental, cu variabile economice; pentru a determina efectele variaţiilor lor, incluzînd cantităţile pe care vrem să le studiem sub formă de parametri, cercetăm cvasiexperimental efectele acestor parametri asupra comportamentului sistemului simbolizat de model.

Funcţia de cvasiexperimentare nu este nouă, ea a fost extinsă în ultimul timp, pe scară foarte largă, îndeosebi în sociologie, prin lucrările lui Simon şi Abelson asupra dinamicii grupurilor umane.

1.9. PERSPECTIVELE APLICAŢIILOR MATEMATICII ÎN DOMENIUL CERCETĂRILOR SOCIOLOGICE

Este deosebit de dificil de a prezenta o imagine clară a metodelor şi a modelelor matematice ce se vor aplica în domeniul sociologiei în viitorii 10, 20, 30 sau chiar 50 de ani. Istoria recentă a acestor aplicaţii pune în evidenţă o serie de discontinuităţi însemnate [10].

Or, ceea ce condiţionează, în final, dezvoltarea unei metodologii eficace de cercetare a fenomenelor şi proceselor sociale sînt îndeosebi evenimentele relativ imprevizibile care au modificat profund cursul cercetării într-un domeniu sau altul, aşa cum au fost lucrările lui C o u r n o t, W a l r a s şi P a r e t o în economie, Claude Lévy S t r a u s s în etnologie, C h o m s k y în lingvistică, ale lui E s t e s şi P i a g e t în psihologie.

În prezent, ceea ce se poate afirma este existenţa anumitor nevoi urgente: dezvoltarea aplicaţiilor analizei relaţiilor cauzale dintre variabilele unei structuri sociale, a modelelor stocastice ale fenomenelor sociale şi a modelelor simulate ale unor procese sociale deosebit de complexe.

Scanned with OKEN Scanner

În mod cert, factorii care vor favoriza extinderea aplicaţiilor matematicii în domeniul sociologiei vor fi :

— perfecţionarea aplicaţiilor tehnicii electronice de calcul în cercetările sociologice ;
— perfecţionarea aplicaţiilor metodelor şi modelelor matematice în cercetările sociologice ;
— creşterea nevoilor pentru ameliorarea calculelor previzionale în sociologie ;
— dezvoltarea şi perfecţionarea învăţămîntului de metodologie, de statistică şi de matematică modernă ;
— interesul crescînd al matematicienilor şi sociologilor în aplicarea matematicii şi a calculatoarelor electronice în ştiinţele umane.

Problemele cele mai abstracte ale teoriei sociologice relevă poate ele înseşi *o metodologie matematică*.

Există toate premisele ca în anii următori să asistăm la o dezvoltare importantă a metodologiei în ştiinţele umane care va cere un aparat matematic din ce în ce mai complex, capabil să reflecte procesele sociale în toată complexitatea lor.

În prezent, există instrumente matematice şi statistice care sînt utilizate direct de toate ştiinţele umane. O serie de preocupări fundamentale sînt comune pentru aceste ştiinţe, cum ar fi : teoria generală a sistemelor, teoria structurală etc.

S-au creat deja o serie de instrumente matematice, aplicabile în ştiinţele umane, îndeosebi în domeniul sociologiei, menţionate mai înainte : metodele statistice-matematice, teoria grafelor, teoria jocurilor, teoria informaţiei sau cibernetica, care au luat un mare avînt în dezvoltarea lor.

Aplicarea metodelor şi modelelor matematice în cercetările sociologice probabil va deschide noi perspective, va crea premisele unei analize cantitative a proceselor sociale fără a se neglija analiza calitativă a lor. La rîndul ei, analiza cantitativă a proceselor sociale va oferi posibilitatea verificării în practică a ipotezelor emise în cercetările sociologice, pentru a face previziuni fundamentate

Scanned with OKEN Scanner

ştiinţific asupra desfăşurării acestor procese, în condiţii diferite. Pe această bază, va fi posibilă — de ce să nu facem o astfel de afirmaţie — o „dirijare" pe baze ştiinţifice a proceselor sociale.

Metodele şi modelele matematice, în continuă dezvoltare şi perfecţionare, probabil că vor contribui într-o măsură importantă la fundamentarea *sociologiei matematice* ca ştiinţă de frontieră, care şi-ar propune drept obiectiv să elaboreze un aparat matematic adecvat analizei, previziunii şi optimizării fenomenelor şi proceselor sociale. Desigur că, extinderea aplicaţiilor matematicii în domeniul sociologiei se va face paralel cu progresele înregistrate atît de matematici cît şi de sociologie. În prezent, nu pot fi ignorate limitele aplicaţiilor matematicii în domeniul sociologiei, legate atît de inconsistenţa instrumentului matematic folosit, cît şi de imperfecţiunea tehnicilor existente de culegere, de prelucrare şi de analiză a informaţiilor sociologice.

Modul de gîndire matematic nu este atotcuprinzător. Tot ceea ce poate fi cuprins cu ajutorul gîndirii matematice rămîne într-o privinţă la limitele unui domeniu restrîns al ştiinţei umane. Domeniul ştiinţelor umane care este însă uriaş, posedă aceleaşi drepturi de a fi luat în considerare ca şi domeniul ştiinţelor naturii, mai exacte.

Unii matematicieni, sînt oarecum dezarmaţi în faţa faptelor sociale. Ei nu pot să înţeleagă aceste fapte ca atare ; nu sînt capabili nici să le explice, „ci doar să stăpînească" un volum mare de informaţie sociologică [11]. De pildă, interdependenţa dintre motivaţia şi acţiunea umană nu poate fi înţeleasă şi explicată de o analiză cantitativă, ci de o analiză calitativă care să elucideze modul de acţiune al oamenilor sub impulsul unui număr mare de motive, care se întrepătrund de cele mai multe ori.

Analiza cantitativă poate cel mult să stăpînească materialul faptic pus la dispoziţie ; prin descompunerea motivelor şi acţiunilor umane în elementele lor componente, analiza logică decelează factorii determinanţi şi relaţiile

matematice esențiale din structurile sociale luate în studiu [12].

Metodele cantitative de analiză a proceselor sociale, prin excelență, fac obiectul statisticii și matematicii clasice. Analizînd latura cantitativă a proceselor sociale, matematica face abstracție de latura lor calitativă, concretă.

Latura cantitativă a realității obiective fiind legată organic de latura calitativă, cunoașterea cantitativă, matematică, a realității sociale poate să ne ducă, în anumite condiții, la concluzii calitative despre această realitate deosebit de complexă.

Într-o strînsă îmbinare cu metodele calitative de cercetare sociologică, metodele cantitative ne dau o imagine mai completă despre procesele sociale studiate. Cu alte cuvinte, aplicarea metodelor cantitative în sociologie este imperativul cunoașterii riguroase, precise, a proceselor sociale.

În condițiile aplicării metodelor cantitative în cercetările sociologice, crește foarte mult rolul analizei calitative a tuturor factorilor care vor fi incluși în modelele matematice ale proceselor sociale [13].

În viitor, eforturile cercetătorilor științifici vor trebui orientate spre elaborarea și perfecționarea metodelor și modelelor matematice adecuate nevoilor de cercetare sociologică, ținîndu-se seama de natura informațiilor obținute din investigațiile de teren și de particularitățile proceselor sociale analizate.

În general procesele sociale sînt deosebit de complexe, avînd un caracter calitativ prin excelență. Pe de altă parte, datorită complexității acestor procese, cuantificarea lor reprezintă un obstacol serios în calea aplicării metodelor cantitative în domeniul sociologiei. Obstacolele generate îndeosebi de cuantificarea proceselor sociale vor putea fi însă învinse, în viitor, prin extinderea aplicațiilor matematicii calitative în acest domeniu, care nu implică cuantificarea acestora. În acest scop, va trebui să aplicăm în viitor cu mai multă hotărîre metodele pe care ni le oferă matematica calitativă, a căror dezvoltare

Scanned with OKEN Scanner

și perfecționare ar putea deschide noi perspective în domeniul cercetărilor sociologice [14].

Utilizarea unor variabile necuantificabile în construirea unor *modele cauzale* ale proceselor sociale va lărgi aria aplicațiilor matematicii în domeniul sociologiei. De pildă, conceptele din teoria mulțimilor și teoria grafelor permit deja folosirea variabilelor calitative în analiza sociologică, fapt care determină creșterea interesului pentru aplicațiile acestor teorii în cercetările sociologice. Prin aplicarea acestor teorii, concomitent cu o serie de elemente ale algebrei matriciale, matematica poate fi aplicată în toate fazele cercetării sociologice, începînd cu emiterea ipotezelor și terminînd cu verificarea lor concretă, ceea ce le conferă o valoare metodologică excepțională pentru cercetările sociologice [15],[16],[17].

Avînd în vedere tendințele existente ale aplicațiilor matematicii în sociologie, întrezărim următoarele direcții principale de dezvoltare și perfecționare în viitor a metodelor matematice aplicate în acest domeniu [9]:

— extinderea aplicațiilor unor metode perfecționate de analiză matematică a structurilor sociale, fundamentate pe teoria mulțimilor, teoria grafelor, teoria corelației și inferențelor cauzale, teoria jocurilor etc.;

— perfecționarea sistemului informațional utilizat în studiul fenomenelor și proceselor sociale, prin extinderea aplicațiilor teoriei informației în cercetările sociologice;

— perfecționarea metodologiei de planificare socială, prin extinderea cercetării operaționale în domeniul sociologiei, care să permită „dirijarea" pe baze științifice a fenomenelor și proceselor sociale, avînd în vedere obiectivele de optimizare urmărite îndeaproape de organele centrale de decizie;

— extinderea aplicațiilor posibile ale calculatoarelor electronice în cercetările sociologice, adică a clasificării automate a informațiilor, a analizei structurilor latente, a simulării proceselor sociale etc.

Scanned with OKEN Scanner

În asemenea condiţii, metodele şi modelele matematice aplicate aduc o contribuţie însemnată la dezvoltarea şi perfecţionarea cercetărilor sociologice concrete.

1.10. CONCLUZII

Extinderea şi perfecţionarea aplicaţiilor, metodelor şi modelelor în cercetările sociologice indică legătura tot mai strînsă care se stabileşte între sociologie şi matematică. Metodele cantitative utilizate în cercetările sociologice se impun tot mai mult în vastul cîmp al problemelor legate de modelarea fenomenelor şi proceselor sociale, ca unul dintre cele mai importante mijloace de cunoaştere ştiinţifică în acest domeniu. Dacă ne gîndim că în domeniul sociologiei raţionăm logic, este util de a da o formă matematică raţionamentului nostru, chiar dacă rezolvăm problema în mod aproximativ. De altfel, dacă noi matematizăm o problemă, nu urmărim altceva, decît de a crea un model aproximativ al realităţii studiate. Un asemenea model poate să nu fie formulat cantitativ, ci calitativ. Se poate afirma că rezolvarea matematică a unui model presupune un spor de informaţie sociologică, deoarece el o face să treacă dintr-o formă implicită pentru noi, într-o formă explicită [16].

Probabil că, în viitor eficacitatea modelelor matematice aplicate în cercetările sociologice va creşte, ca urmare a construirii unui complex de modele calitative şi cantitative ale proceselor sociale, organizării unui sistem unic de informaţii sociale şi a creării unei reţele unice de stat de centre de calcul care să asigure prelucrarea şi interpretarea acestor informaţii [17], [18], [19]. În asemenea condiţii, metodele şi modelele matematice elaborate vor constitui un cîştig pentru cercetările sociologice concrete, îndeosebi pentru cele cu caracter metodologic.

Scanned with OKEN Scanner

BIBLIOGRAFIE

1. R. Hayward şi Alker Jr., *Mathematics & Politics*, New York, The Macmillan Company, London, Collier — Macmillan — Limited, 1965.
2. Kenneth H. Arrow, *Mathematical models in the social sciences*, în: Lerner Daniel şi Lass Well Harold D., *The policy sciences*, Stanford, California, Stanford University Press, 1949.
3. Henri H. Stahl, *Păreri cu privire la actuala răscruce a sociologiei*, Sociologia Militans, I, Obiectul sociologiei, Bucureşti, Editura ştiinţifică p. 44—58.
4. Raymond Boudon, *L'analyse mathématique des faits sociaux*, Paris, Plon, 1967.
5. James S. Coleman, *Introduction to mathematical sociology*, Glencoe, The Free Press, 1964.
6. Raymond Boudon, *Méthodes d'analyse causale*, în: Revue française de sociologie, VI, 1965.
7. Raymond Boudon, *La syntèse de la sociologie, Introduction à l'analyse mathématique des faits sociaux*, Paris, Plon, 1966.
8. Simon Herbert, *Models of man, social and rational, Mathematical essays on rational behavior in social setting*, New York, 1957.
9. Tiberiu Schatteles, *Modelul în ştiinţele economice*, Bucureşti, Edit. politică, 1967.
10. Raymond Boudon, *L'étude internationale concernant les tendences d'application des Mathématiques dans les sciences sociales*, Paris, UNESCO, 1967.
11. C. A. Moser, *Metodele de anchetă în investigarea fenomenelor sociale*, Bucureşti, Edit. ştiinţifică, 1967.
12. Raymond Boudon et Paul F. Lazarsfeld, *Le vocabulaire des sciences sociales*, Paris, Le Haye, Mouton, 1965.
13. Claude Flament, *Théorie des graphes et structures sociales*, Paris, Mouton, Gauthier — Villars, 1965.
14. Martin Shubik, *Games theory and related approaches to social behavior*, New York, Wiley, 1964.

Scanned with OKEN Scanner

15. I. A. Levada, *Metodele exacte în cercetarea sociologică*, în: *Teorie și metodă în științele sociale*, vol. II, București, Edit. politică, 1966.
16. Claude Berge, *Théorie des graphes et ses applications*, Paris, Dunod, 1958.
17. A. I. Khinchin, *Mathematical foundation of information theory*, New York, Dover, 1957.
18. Н. Н. Ряузов ș. a., *Статистика*, Москва, Изд. «Статистика», p. 30—110.
19. Vitali Stoljarov, *Zur Technik und Methodologie einiger quantifizierender Methoden der soziologischen Forschung*, Berlin, Dietz, 1966.

Scanned with OKEN Scanner

2. MODELE INDUCTIVE

2.1. INTRODUCERE

În ultimii ani, asistăm la o extindere din ce în ce mai mare a aplicaţiilor teoriei grafelor şi a algebrei matriciale în domeniul sociologiei, ştiinţă care acordă o importanţă crescîndă studiului proceselor sociale. De altfel, sociologii se numără printre primii care au folosit teoria grafelor în activitatea de investigaţie ştiinţifică, tocmai datorită conţinutului extrem de bogat al acestei ramuri a teoriei mulţimilor.

Într-adevăr printr-un graf şi o matrice asociată lui pot fi reprezentate procese sociale foarte diferite, ca de exemplu : procesul de mobilitate socială şi profesională, procesul de migraţie a populaţiei, procesul de dezvoltare urbană, procesul de industrializare şi de transformare cooperatistă a agriculturii etc.

Teoria grafelor şi calculul matricial pot lărgi domeniul de cunoaştere al proceselor sociale, deoarece implică elaborarea unor construcţii logice, simbolice şi intuitive ale structurilor sociale, aflate într-o continuă schimbare şi transformare. De pildă, procesul de urbanizare poate fi privit ca un proces de restructurare socială, care antrenează mişcări de populaţie între unele ramuri ale producţiei materiale, regrupări sociale şi profesionale, fluxuri migratorii teritoriale etc. [1].

În domeniul sociologiei, teoria grafelor şi calculul matricial pot aduce, în viitor, un aport însemnat în formularea unor modele inductive de relaţii cauzale, adică în construcţia unor asemenea reprezentări care permit analiza şi interpretarea pe baze ştiinţifice a structurilor

Scanned with OKEN Scanner

și proceselor sociale deosebit de complexe. Asemenea reprezentări ar putea contribui la formularea de ipoteze verificabile asupra desfășurării proceselor sociale, cunoscînd faptul că *un model inductiv este de obicei un ansamblu unitar de ipoteze de lucru* [2].

Pentru fiecare proces social luat în studiu, cercetătorul științific trebuie să stabilească un model inductiv, matematizabil sau nu, adică un grup complex de ipoteze de lucru, emise într-un proces de gîndire care apelează la tehnicile raționamentului experimental sau logic. Elaborarea unui asemenea model implică alegerea, din masa complexă a realității sociale, a unor fapte, aspecte și observații, cărora din punct de vedere teoretic le acordăm o semnificație deosebită [2].

În cercetarea unui proces social dat trebuie să analizăm realitatea socială luată în studiu, de fapt exterioară nouă, avînd în vedere condițiile geografice, aspectele demografice, psihologia grupurilor umane, istoria imediat trecută a fenomenului social, aspectele culturale, economice, juridice și politice ale societății [3].

În faza de elaborare a ipotezelor de lucru vom apela la *metoda inductivă* de studiu a proceselor sociale aplicată prin următoarele etape principale [4]:

a) formularea de previziuni asupra relațiilor de dependență între indicatorii aleși; cercetătorul poate emite o anumită supoziție asupra existenței, în anumite condiții date, a unor relații de dependență, bazîndu-se pe experiența proprie, pe informațiile acumulate în activitatea de cercetare științifică și pe raționamente logice;

b) efectuarea unui sondaj pilot prin care se pun în evidență relațiile de dependență dintre indicatorii selectați;

c) emiterea ipotezei verificabile pe baza analizei informațiilor culese cu ocazia efectuării anchetei-pilot. În această etapă, cercetătorul confruntă informațiile obținute din cercetarea-pilot cu supoziția sa asupra relațiilor de dependență dintre indicatorii analizați. Dacă în toate cazurile, fără excepție, aceste informații concordă

Scanned with OKEN Scanner

cu previziunile făcute, atunci acestea din urmă se transformă în ipoteze. În caz contrar, previziunile făcute se pun de acord cu informaţiile obţinute din cercetarea-pilot.

O concluzie inductivă în domeniul cercetărilor sociologice nu este totdeauna justă, datorită faptului că în cercetările respective pot fi omise unele observaţii sau sondaje-pilot strict necesare, fapt care poate denatura ipotezele emise asupra relaţiilor de dependenţă. De aceea, cu ajutorul metodei inductive aplicate în cercetările sociologice putem doar să intuim, din observaţii particulare, legi cu o valabilitate generală.

2.2 CONCEPTELE DE PROCES, STRUCTURĂ ŞI MODEL

Modelarea proceselor sociale oferă posibilităţi de experimentare pentru cunoaşterea tendinţelor de dezvoltare socială în variantele cele mai probabile. De pildă, anumite procese sociale particulare, cum ar fi procesul de migraţie a populaţiei de la sat la oraş, procesul de transmitere a culturii în mediul rural, procesul de mobilitate şi stratificare socială etc., pot fi modelate pornindu-se de la datele statistice privind condiţiile producerii lor şi de la diferite aspecte ale evoluţiei trecute, în scopul cunoaşterii dinamicii lor în variantele cele mai probabile de realizare. Modelul permite să se determine, ca previziuni, caracteristicile evoluţiei procesului social luat în studiu. În asemenea condiţii, modelul unui proces social dat — pe baza unor informaţii observabile noi — poate fi refăcut şi verificat, asigurîndu-se în acest mod perfecţionarea lui, fără să se aştepte evoluţia efectivă a procesului analizat.

Necesităţile de ameliorare a calculelor previzionale a determinat pe oamenii de ştiinţă să extindă şi să perfecţioneze *metodele de simulare a dezvoltării sociale.*

Simularea dezvoltării sociale reprezintă un ansamblu de procedee de calcul, care aparţin de regulă matematicii

moderne și tehnicii electronice, procedee folosite în studiul flexibilității sistemelor sociale la modificările efectelor factorilor de influență, în scopul determinării comportamentului lor în viitor, în variantele cele mai probabile de realizare.

Prin urmare, metoda simulării dezvoltării sociale apelează la conceptul de *sistem social*, care reprezintă unitatea dialectică dintre *structura socială* și *procesul social*. Într-un sistem social dat, structura socială conferă sistemului proprietatea de cuprindere a elementelor interconectate într-un tot unitar, iar procesul social reflectă capacitatea de dezvoltare a sistemului în timp.

Structura socială poate fi interpretată ca un sistem de colectivități umane avînd caracteristici diferite, reunite în așa fel încît să constituie un organism social complex, cuprinzînd un ansamblu de relații între elementele structurii.

Procesul social poate fi privit ca o desfășurare în timp a unui complex de fenomene și fapte sociale.

Spre deosebire de structura socială, care reflectă latura statică a realității sociale luate în studiu, procesul social reprezintă conținutul dinamic și evolutiv al acestei realități.

Schema de modelare sociologică a proceselor sociale — unele matematizabile, altele nu — prezintă un interes deosebit în practica cercetărilor sociologice concrete. Asemenea schemă presupune o sinteză a proceselor particulare care determină evoluția acestor procese deosebit de generale. De pildă, în analiza structurilor sociale concrete putem apela la scheme operaționale proprii de investigație sociologică, care reunesc toate științele sociale particulare : economice, juridice, culturale, politice, geografice, demografice, istorice și psihologice.

Structurile sociale analizate cu ajutorul schemelor de modelare sociologică nu pot fi înțelese decît prin analiza exhaustivă a elementelor lor componente, importanța fiecăruia dintre ele depinzînd îndeosebi de locul, ponderea și rolul funcțional pe care îl au în structura luată în studiu.

Scanned with OKEN Scanner

O schemă-model implică o cunoaştere aprofundată a structurii şi procesului social luat în studiu şi o concepţie teoretică de bază.

Conceptul de model s-a dezvoltat continuu, în funcţie de evoluţia ideilor despre structura socială şi procesul social.

Este logic să ne punem întrebarea : ce înţelegem prin model ? *Modelul unui proces social dat reprezintă o schemă simbolică, simplificată şi aproximativă, formulată sau nu în limbaj matematic şi destinată explicării modului de desfăşurare în timp a unei acţiuni umane.*

Modelul unui asemenea proces poate fi privit ca un sistem unitar de variabile (implicite sau explicite) şi de relaţii de dependenţă dintre ele (formulate sau nu în limbaj matematic) destinat analizei aspectelor dinamice ale acţiunii umane în timp.

Orice model al unui proces social dat poate fi caracterizat prin următoarele elemente de bază :

a) *variabila* — un element definitiv şi revelator al unui factor particular de natură diferită : social-cultural, economic, biologic, psihic, istoric, politic, juridic etc. ;

b) *relaţia de dependenţă* — raportul între două variabile (implicite sau explicite), din care una este *cauza* şi cealaltă este *efectul*.

Mulţimea variabilelor explicite sau implicite şi a relaţiilor de dependenţă servesc elaborării *modelelor matematice* ale proceselor sociale luate în studiu.

Mulţimea finită a metodelor matematice ale proceselor sociale (M) poate fi caracterizată de mulţimea finită a variabilelor (X) şi de mulţimea relaţiilor de dependenţă dintre ele (Y), care poate fi descrisă cu ajutorul unor elemente din teoria mulţimilor, în felul următor :

$\{M\} = \{X, Y\}$, unde X reprezintă variabile explicite sau implicite şi Y relaţii de dependenţă.

În structurile proceselor sociale deosebit de complexe există un sistem de relaţii de dependenţă între indica-

torii care reflectă o serie de factori de natură diferită : economică, culturală, geografică, biologică, istorică, psihică, politică, juridică etc. Relaţiile de dependenţă existente într-o astfel de structură pot fi esenţiale sau neesenţiale, importante sau mai puţin importante.

Avînd în vedere complexitatea structurilor sociale, de cele mai multe ori sîntem tentaţi să studiem toate relaţiile de dependenţă existente. Aceasta este, fără îndoială, foarte just, dar nu mai puţin adevărat este şi faptul că nu toate relaţiile de dependenţă dintre indicatorii care caracterizează o structură socială dată au acelaşi grad de tărie, aceeaşi importanţă. De aceea, trebuie să ţinem cont nu numai de existenţa relaţiilor de dependenţă, ci şi de faptul dacă ele sînt esenţiale sau neesenţiale pentru structura socială analizată. Cînd nu există o viziune clară asupra proceselor sociale studiate, nu este posibilă deosebirea esenţialului de neesenţial, ca urmare a incapacităţii de abstractizare. De aceea, în cadrul cercetărilor sociologice trebuie să ştim ce fel de relaţii de dependenţă trebuie luate neapărat în considerare şi care sînt relaţiile de care trebuie să facem abstracţie. Fără această abstracţie este imposibilă descoperirea unor legi generale.

Cînd analizăm relaţiile de dependenţă din structura unui proces social oarecare, trebuie să examinăm relaţiile esenţiale dintre variabilele care alcătuiesc un sistem, făcînd abstracţie de relaţiile neesenţiale, întîmplătoare. Descoperirea relaţiilor de dependenţă esenţiale este importantă pentru cunoaşterea esenţei proceselor sociale şi, invers, dezvăluirea esenţei acestor procese este necesară pentru cunoaşterea relaţiilor de dependenţă esenţiale din structura socială studiată.

Cunoaşterea tuturor relaţiilor de dependenţă dintre variabilele care caracterizează o structură socială ne procură o serie de informaţii abstracte, dar sub raport gnoseologic spun prea puţin despre structura analizată.

Cunoscînd natura relaţiilor de dependenţă — esenţiale sau neesenţiale — ştim incomparabil mai mult, putem descoperi astfel esenţa procesului social analizat

Scanned with OKEN Scanner

De aceea, construcţiile intuitive ale structurii unor procese sociale, cercetate cu ajutorul grafelor şi matricilor asociate lor, permit punerea în evidenţă a relaţiilor de dependenţă dintre variabile şi a factorilor determinaţi şi determinanţi din structura respectivă, concepută ca un sistem unitar.

Analizînd relaţiile de dependenţă dintre variabile, nu trebuie să facem o diferenţiere rigidă, nedialectică, între *cauză* şi *efect*, ignorînd acţiunile reciproce. În acelaşi timp, trebuie să ţinem seama şi de faptul că relaţiile de dependenţă dintre variabile nu sînt date o dată pentru totdeauna şi nu pot fi examinate doar în mod static. Dacă procesele sociale particulare din structura unui proces social general, deosebit de complex, se dezvoltă continuu, rezultă că şi relaţiile de dependenţă pe care le angrenează în structura analizată, se schimbă mereu.

Prin descifrarea sistemului de relaţii de dependenţă dintre variabilele unei structuri sociale, evidenţiem unitatea concretă a procesului social analizat. Desigur că, descifrarea unui sistem unitar de relaţii de dependenţă este util să se facă cu ajutorul matematicii calitative. De pildă, prin aplicarea unor concepte din teoria mulţimilor, din teoria grafelor şi din algebra matricială, putem efectua analiza matematică calitativă a structurii unui proces social oarecare. Aplicarea acestor teorii matematice în studiul proceselor sociale prezintă o valoare metodologică excepţională pentru cercetările sociologice concrete.

Între conceptele de *relaţie de dependenţă* şi de *relaţie cauzală* există diferenţe de nuanţă.

Ce este cauzalitatea?

Cauzalitatea este o categorie filozofică care exprimă interacţiunea dintre *cauză* şi *efect*, constituind una dintre principalele forme ale conexiunii fenomenelor şi proceselor sociale. Această categorie stă la baza concepţiei deterministe sau a determinismului.

Cauzalitatea are un caracter obiectiv. Orice fenomen sau proces social are o cauză; nu există fenomene sau procese sociale fără cauză.

Scanned with OKEN Scanner

Conceptul de *relaţie cauzală* exprimă un raport între o acţiune sau inacţiune şi urmările imediate ale acesteia, raport în care acţiunea sau inacţiunea precedă sau determină în mod obiectiv aceste urmări [1].

Într-o relaţie cauzală trebuie să definim categoriile de cauză şi efect, care desemnează una dintre cele mai generale forme ale conexiunii fenomenelor şi proceselor sociale [2].

Cauza este un fenomen sau complex de fenomene, care precedă şi produce (provoacă, generează sau determină) un alt fenomen, denumit efect. Acţiunea unei cauze se manifestă întotdeauna în anumite condiţii, influenţînd corelaţia dintre cauză şi efect.

Raporturile cauzale au un caracter obiectiv, putînd fi verificate în activitatea practică. Aceste raporturi pot fi de diferite tipuri, iar cauzele pot fi multiple (principale şi secundare, esenţiale şi neesenţiale, necesare şi întîmplătoare, interne şi externe).

Efectul este un fenomen care urmează unui alt fenomen, numit cauză şi produs în mod necesar de acesta.

Prin *inferenţă cauzală* vom înţelege operaţia logică de la cauză la efect, în care ultima noţiune este dedusă din prima.

Relaţiile dintre variabilele unei structuri sociale oarecare sînt caracterizate de următorii indicatori [5]:

— indicatori corelaţionali;
— indicatori definitorii;
— indicatori inferenţiali (deductivi).

Indicatorii corelaţionali exprimă o variaţie concomitentă — a două sau mai multe variabile — care poate fi observată în mod direct.

Indicatorii definitorii descriu o relaţie dintre două sau mai multe variabile, evident pe baza unei convenţii terminologice.

---

[1] Noţiunea de relaţie cauzală este sinonimă cu cea de legătură de cauzalitate, aşa cum este definită în *Dicţionarul enciclopedic*, vol. 2, p. 215.

[2] Noţiunile de *cauză* şi *efect* sînt luate în înţelesul din *Dicţionarul enciclopedic*, vol. 2, p. 215.

Scanned with OKEN Scanner

Indicatorii inferențiali indică relațiile de dependență care nu sînt direct observabile, dar sînt deduse din modelul raporturilor dintre variabilele observate, în scopul explicării acestor raporturi.

Analiza matematică a acestor grupe de indicatori poate fi extrem de utilă în studiul structurilor sociale.

2.3. UNELE APLICAȚII GENERALE ALE TEORIEI MULȚIMILOR ȘI ALE TEORIEI GRAFELOR ÎN ANALIZA PROCESELOR SOCIALE

Sistemul compus din elemente interdependente și proprietăți ale ansamblului dat rezultă din unitatea subansamblurilor și elementelor componente ale acestora. De pildă, dacă apelăm la noțiuni din teoria mulțimilor, ansamblul A construit din reunirea subansamblurilor interdependente S_k se poate nota prin:

$$A = \bigcup_{k=1}^{r} S_k$$

Aici ansamblul A poate reprezenta un proces social oarecare (proces general), iar subansamblurile sale S_k pot fi procese particulare ce caracterizează structural procesul general. Se pot da exemple numeroase de procese sociale complexe, cum ar fi: procesul de migrație a populației de la sat la oraș, procesul de transmitere a culturii, procesul de mobilitate și stratificare socială etc. caracterizate structural de o serie de procese particulare interdependente, derivînd din grupul de procese economice, culturale, istorice, geografice, psihice, juridice, politice etc. (k). Unitatea acestor procese particulare constituie procesul social în toată complexitatea lui.

Una din noțiunile de bază ale matematicii moderne este aceea de *mulțime* sau de *ansamblu*. Un ansamblu este

alcătuit dintr-un număr finit de subansambluri, formate dintr-o serie de elemente. De exemplu, pentru un proces social oarecare, componentele ansamblului pot fi indicatori și relații de dependență, care vor caracteriza unele procese particulare (ansamble ale procesului general). Se pot da exemple numeroase de indicatori cum sînt: gradul de școlarizare, gradul de calificare, structura de vîrstă, nivelul de salarizare etc. Dacă vom nota indicatorii cu $x_j^{(k)}$ iar relațiile de dependență dintre acești indicatori cu y_{ij}, vom scrie:

$$A = \bigcup_{k=1}^{r} S_k = \{x_j^{(k)};\ y_{ij} \mid x_j^{(k)} \in S_k \in A \text{ și } Y_{ij} \in A\}$$

Structura unui astfel de ansamblu care caracterizează un proces social complex este reprezentată în figura 5 și poate fi interpretată în felul următor: procesul social

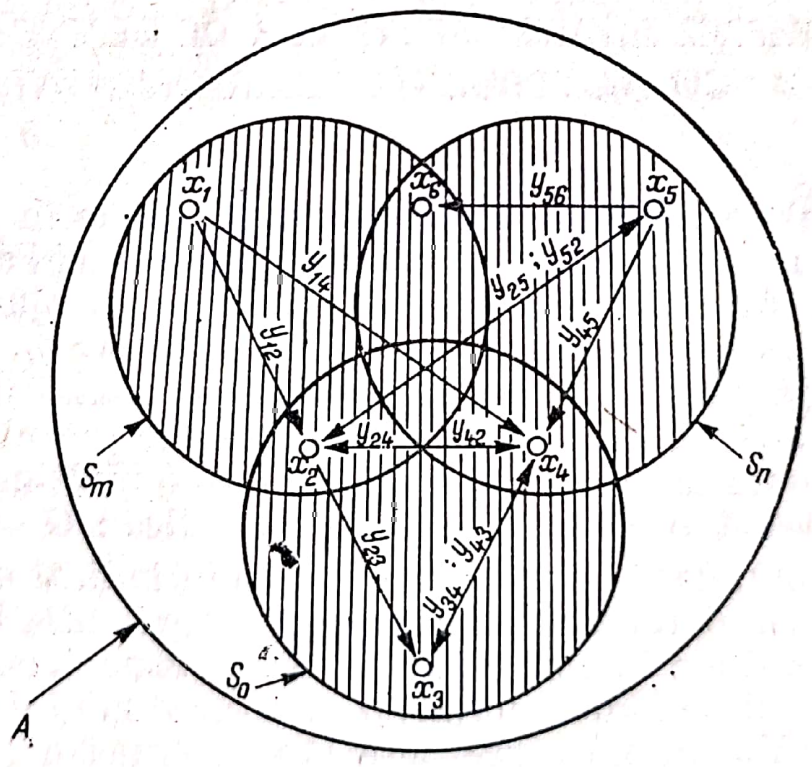


Fig. 5. — Modelul de dependență al procesului social

general A constituit din procesele particulare interdependente S_k este caracterizat structural de o mulțime finită de indicatori ai proceselor particulare $x_j^{(k)}$ și de relații de dependență y_{ij} dintre acești indicatori. În acest caz, între mulțimea indicatorilor și mulțimea relațiilor de dependență se stabilește o corespondență care stă la baza formulării unui model de dependență. Într-un astfel de model, elementele pot fi reprezentate grafic, în așa fel încît mulțimea variabilelor $x_j^{(k)}$ să constituie mulțimea nodurilor X, iar mulțimea relațiilor cauzale y_{ij} să formeze mulțimea arcelor orientate Y. Ansamblul acestor variabile și relații va constitui un model cauzal $G = (X, Y)$, care va purta denumirea de graf orientat. Grafele au o bază matematică, deoarece teoria acestora constituie una din ramurile teoriei mulțimilor, deosebit de utilă în rezolvarea problemelor în care intervin combinări sau ordonări de elemente în structura unui ansamblu oarecare.

2.4. REPREZENTAREA RELAȚIILOR DE DEPENDENȚĂ CU AJUTORUL GRAFELOR ORIENTATE

Graful se prezintă sub forma unui grafic în care segmentele, neorientate sau orientate, leagă perechi de puncte (noduri), formînd arce ce pot fi notate cu ajutorul elementelor x_i și x_j pe care le unește. La orice arc $y_{ij} = (x_i, x_j)$ se distinge punctul x_i sau extremitatea inițială și punctul x_j sau extremitatea finală. Prin urmare, graful poate fi definit de un ansamblu de puncte X și o mulțime de arce Y care leagă o serie de noduri x_j, adică: $G = (X, Y)$. De exemplu, dacă punctele x_j reprezintă indicatorii sociali ai progresului tehnic, iar arcele y_{ij} reprezintă legăturile cauzale posibile între acești indicatori, atunci, cu ajutorul ansamblului de indicatori (X) și al ansamblului de legături posibile (Y), putem reprezenta graful structurii relațiilor de dependență, graf care caracterizează progresul tehnic.

Scanned with OKEN Scanner

Pentru aceasta, să admitem că ansamblul X este caracterizat de mulțimea x_j a indicatorilor din tabelul *1*.

Tabelul 1

Nr. crt.	Denumirea indicatorului	Simbolul
1	Nivelul de tehnicitate	x_1
2	Categoria de încadrare a muncitorului	x_2
3	Vechimea în cîmpul muncii	x_3
4	Nivelul de cultură al muncitorului	x_4
5	Gradul de calificare	x_5
6	Vîrsta	x_6
7	Apartenența politică	x_7

Considerăm că între indicatorii x_j $(j = 1, 2, \ldots, n)$ există următoarele corelații statistice: (x_1, x_2), (x_1, x_4), (x_1, x_5), (x_2, x_3), (x_3, x_5), (x_2, x_5), (x_4, x_5), (x_2, x_5), (x_1, x_7), (x_4, x_7) și (x_5, x_7). Să presupunem că aceste corelații reprezintă arce neorientate în graficul rețea.

Pe baza indicatorilor (X) și a corelațiilor dintre acești indicatori (Y), putem construi graful neorientat $G = (X, Y)$ din fig. 6.

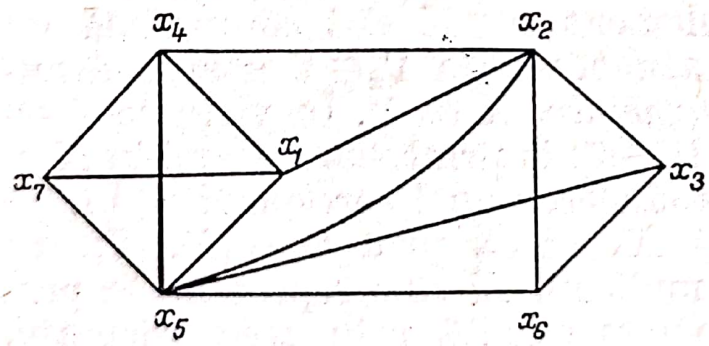


Fig. 6. — Graf neorientat

Graficele rețea elaborate pentru un ansamblu de relații de dependență au diferite forme care sînt supuse transformărilor geometrice, conservînd însă proprietatea de a reprezenta configurația relațiilor existente într-o structură socială dată. Astfel, lungimile arcelor și unghiu-

rile dintre ele pot fi modificate fără ca prin aceasta să se schimbe gradul de reprezentativitate al graficului rețea. Această proprietate este folosită la elaborarea graficelor rețea cît mai compacte și cît mai clare, astfel încît reprezentarea relațiilor de dependență dintre variabile să fie cît mai intuitivă.

Graful poate fi orientat sau neorientat după cum arcele sale au sau nu un sens indicat prin săgeți. În acest caz, arcele neorientate marchează existența unor corelații statistice, pe cînd arcele orientate pun în evidență relații de dependență (sau relații cauzale). De aceea, grafurile neorientate se folosesc atunci cînd putem evidenția existența unor asocieri dintre variabile, pe cînd grafurile orientate — atunci cînd putem sesiza efectele unor variabile asupra altora.

Graful orientat are o valoare cognitivă mult mai mare decît graful neorientat (fig. 6).

În studiul grafelor orientate intervin următoarele elemente :

a) mulțimea punctelor care se notează cu X și ale cărei elemente sînt $x_j \in X$; $i \in N = \{0, 1, 2, 3, \ldots, n\}$;

b) mulțimea arcelor care se notează cu Y și ale cărei elemente sînt $y_{ij} \in Y$; $i, j = N$; $i \neq j$.

Dacă sînt date : două ansambluri X și Y și o lege în virtutea căreia oricărui element $x_j \in X$ să-i corespundă un ansamblu bine definit, atunci relația $\Gamma_x \in Y$ poartă denumirea de *aplicație a ansamblului X în Y*. În virtutea acestui fapt, un graf poate fi definit printr-un ansamblu X și o lege care permite introducerea unei corespondențe între elementele ansamblului X, adică prin $G = (X, \Gamma)$. În acest graf, elementele mulțimii X sînt reprezentate prin puncte, iar relațiile de dependență prin arce orientate, care leagă fiecare element $x_j \in X$ cu elementele subansamblului $\Gamma_x \in X$. Spre exemplificare, pentru ansamblul de variabile $X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$ din fig. 8, vom admite că aplicația Γ_x este caracterizată prin :

$$\Gamma x_1 = \{x_2, x_4, x_5\} ;$$

$$\Gamma x_2 = \{0\};$$

$$\Gamma x_3 = \{x_2, x_5\};$$

$$\Gamma x_4 = \{x_2, x_5\};$$

$$\Gamma x_5 = \{x_2\};$$

$$\Gamma x_6 = \{x_2, x_3, x_5\};$$

$$\Gamma x_7 = \{x_1, x_4, x_5\}.$$

Pe baza ansamblului X de indicatori adaptat şi a aplicaţiei Γx de mai sus, putem construi graful $G = (X, \Gamma)$ în următoarea succesiune: mai întîi se marchează prin

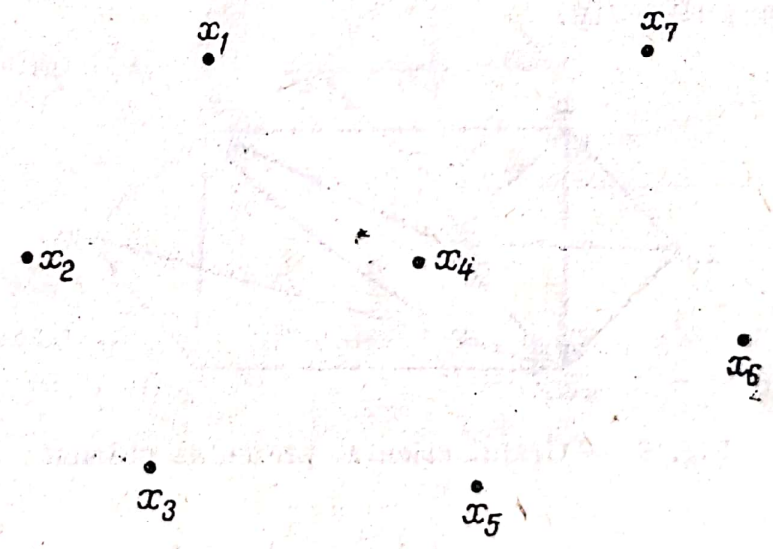


Fig. 7. — Reprezentarea nodurilor (punctelor) grafului

noduri sau puncte indicatorii x_j ai ansamblului X (fig. 7); după aceea se trasează arcele (drepte sau curbe) care leagă nodurile (punctele), avînd în vedere aplicaţia Γx speci-

ficată mai înainte (fig. 8), şi în final se transformă graful obţinut într-unul cît mai compact şi cît mai clar (fig. 9).

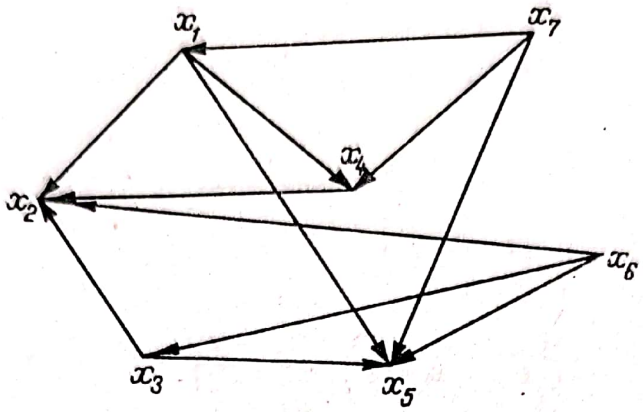


Fig. 8. — Reprezentarea grafului orientat

Un graf, fie el orientat sau neorientat, este alcătuit din mai multe rețele, care pot forma grafe reduse, grafe parțiale și subgrafe.

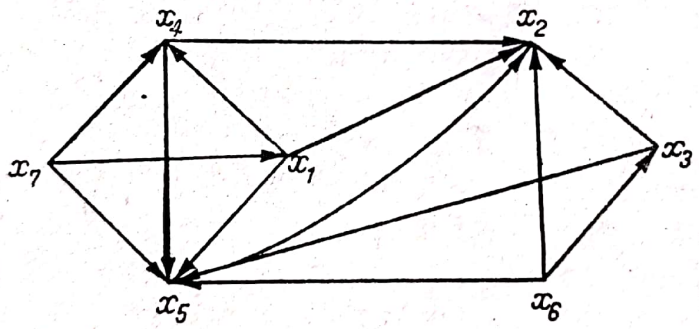


Fig. 9. — Graful orientat prezentat ordonat

Graful redus este graful obţinut dintr-un graf obişnuit în care unele noduri s-au grupat într-unul singur. Astfel, în graful din fig 10, o serie de puncte pot fi grupate ca în fig. 6.

Graful parțial este graful obținut dintr-un graf obişnuit în care s-au suprimat unul sau mai multe arce, numărul nodurilor rămînînd acelaşi. De exemplu, prin

suprimarea arcelor (x_4, x_2) şi (x_5, x_2) din graful prezentat în fig. 5, obţinem graful parţial din fig. 11.

Subgraful se obţine dintr-un graf obişnuit în care s-au suprimat unul sau mai multe noduri şi toate arcele ce

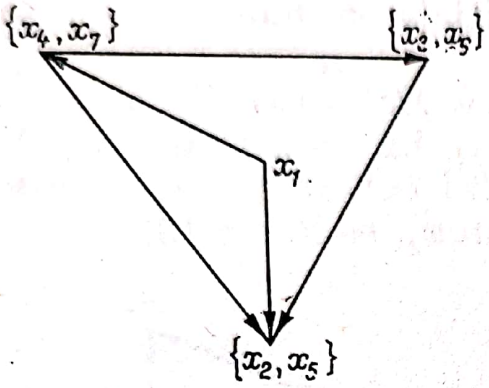
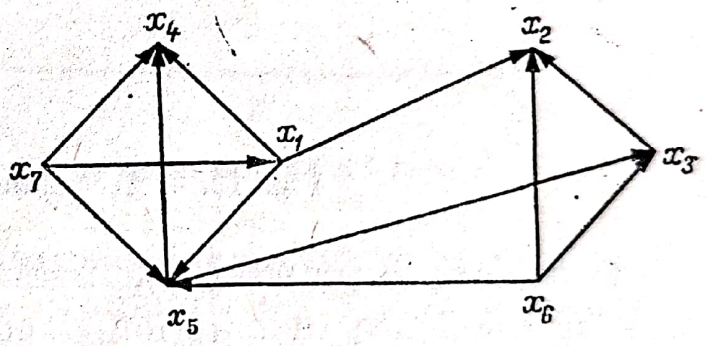


Fig. 10. — Graful redus

Fig. 11. — Graful parţial



leagă aceste noduri cu celelalte noduri ale grafului. De exemplu, prin suprimarea punctului x_3 din graful prezentat în fig. 5, obţinem subgraful din fig. 12.

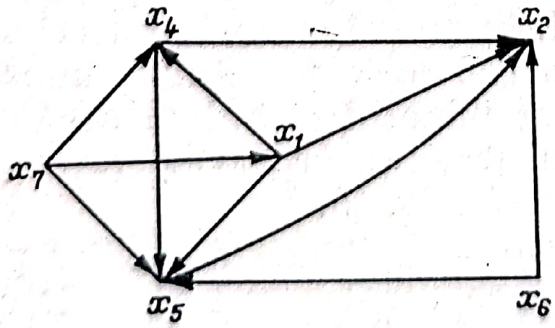


Fig. 12. — Subgraful

2.5. ELEMENTE DE BAZĂ ALE GRAFULUI ORIENTAT

Un graf este alcătuit din noduri şi arce orientate care constituie elementele de bază ale acestuia. La 2.4 am văzut că la orice arc orientat $y_{ij} = (x_1, x_3)$ putem distinge punctul x_i sau extremitatea iniţială şi punctul x_j sau extremitatea finală.

Două arce sînt adiacente dacă sînt distincte şi au o extremitate comună. De exemplu, în graful din fig. 5, arcele (x_1, x_2) şi (x_1, x_5) au nodul comun x_1, iar arcele (x_1, x_2) şi (x_6, x_2) au punctul comun x_2, ca în fig. 13.

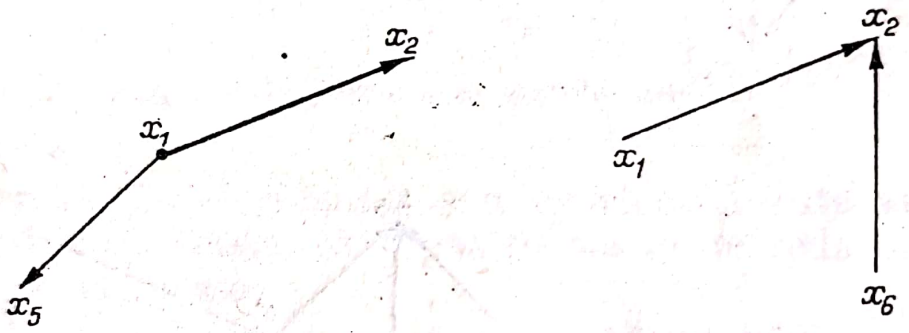


Fig. 13. — Arce adiacente

Considerînd punctul x_i, notăm cu $d^+_{(x_i)}$ ansamblul arcelor incidente în exterior şi cu $d^-_{(x_i)}$ ansamblul arcelor incidente în interior. De exemplu, pentru nodul x_4 din graful prezentat în fig. 9, ansamblul arcelor incidente în interior s-a dat în fig. 14, iar al arcelor incidente în exterior în fig. 15.

În teoria grafelor se utilizează şi noţiunea de *drum* care reprezintă o succesiune de arce adiacente, care permite trecerea de la un nod la altul, în sensul orientării arcelor, adică o succesiune de arce, astfel încît punctul terminal al unui arc să corespundă cu punctul iniţial al arcului următor. De exemplu, drumul (x_7, x_1), (x_1, x_5) şi (x_5, x_3) din fig. 9 este reprezentat, izolat, în fig. 16.

Un drum se determină prin succesiunea nodurilor pe care le parcurge sau succesiunea arcelor ce-l alcătuiesc.

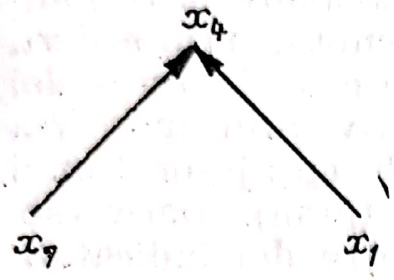


Fig. 14. — Ansamblu arcelor incidente în interior

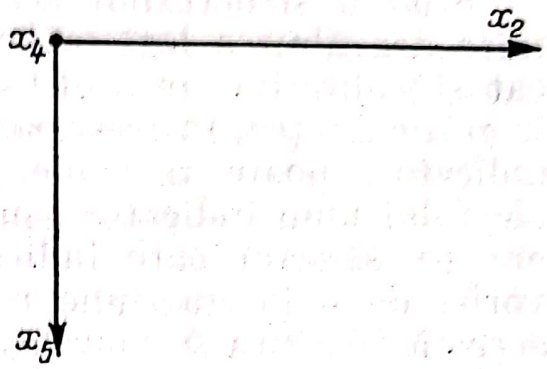


Fig. 15. — Ansamblul arcelor incidente în exterior

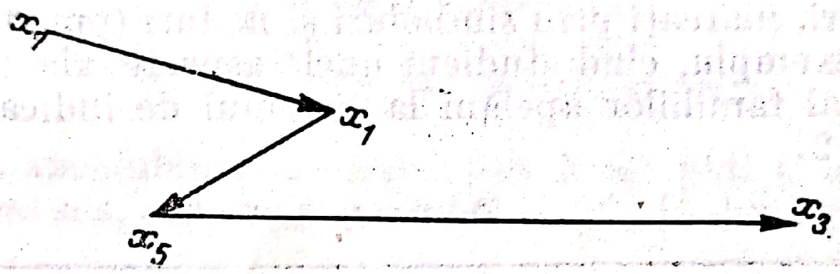


Fig. 16. — Exemplu de drum în graficul rețea

Lungimea unui drum este determinată de numărul arcelor care intră în alcătuirea drumului. De exemplu, drumul (x_7, x_1), (x_1, x_5) și (x_5, x_3) din graful prezentat în fig. 5, are o lungime de trei arce.

Un drum pentru care punctul său inițial corespunde cu punctul final constituie un *circuit*.

2.6. GRAFUL ȘI DRUMURILE ÎNTR-UN GRAF ORIENTAT

O structură socială alcătuită din relații de dependență poate fi reprezentată cu ajutorul unui *graf orientat*. Elementele acestei structuri sînt indicatori sau variabile, iar ansamblul de variabile și relații dintre ele cu ajutorul cărora poate fi analizat fenomenul sau procesul social luat în studiu constituie *sistemul social*.

Scanned with OKEN Scanner

Studiul structurilor sociale cu ajutorul grafelor permite dezvăluirea legăturilor cauzale existente între indicatori (calitativi sau cantitativi) reprezentanţi prin *noduri*, în graficul reţea. În acest caz, o legătură cauzală dintre doi indicatori poate fi evidenţiată sugestiv prin marcarea efectului unui indicator asupra celuilalt cu ajutorul unui *arc* (o săgeată care indică sensul acţiunii). Dacă este vorba de o interacţiune reciprocă dintre doi indicatori, această legătură se marchează în graficul reţea cu două arce avînd sensuri opuse sau cu un singur arc avînd două săgeţi opuse la extremităţi.

În virtutea celor arătate mai înainte, graful reprezintă un ansamblu de cupluri (x_i, x_j) alcătuite din perechi de indicatori, marcaţi prin simboluri şi noduri (puncte).

De exemplu, cînd studiem unele aspecte ale nivelului de trai al familiilor apelăm la sistemul de indicatori din tabelul 2.

Tabelul 2

Nr. crt.	Denumirea indicatorului	Simbolul
1	Numărul membrilor în familie	x_1
2	Numărul membrilor de familie întreţinuţi	x_2
3	Numărul copiilor de vîrstă preşcolară	x_3
4	Numărul persoanelor ocupate în cîmpul muncii	x_4
5	Salariul total al persoanelor aflate în cîmpul muncii	x_5
6	Mărimea suprafeţei locuibile ocupată de familie	x_6
7	Nivelul de dezvoltare culturală al membrilor familiei	x_7
8	Sporul la învăţătură al copiilor	x_8
9	Veniturile familiei, cu excepţia salariilor de bază	x_9
10	Costul minim de trai al familiei	x_{10}

Să admitem că pentru ansamblul acestor indicatori, pe care îi notăm cu $X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$, aplicaţia Γx este caracterizată prin :

$$\Gamma x_1 = \{x_4, x_6, x_9, x_{10}\};$$

$$\Gamma x_2 = \{x_1, x_4, x_6, x_9\};$$
$$\Gamma x_3 = \{x_1, x_2\ x_{10}\};$$
$$\Gamma x_4 = \{x_5, x_7\};$$
$$\Gamma x_5 = \{x_2, x_3, x_4, x_7, x_9\};$$
$$\Gamma x_6 = \{x_1, x_2, x_3, x_7, x_8\};$$
$$\Gamma x_7 = \{x_5, x_8. x_9\};$$
$$\Gamma x_8 = \{0\};$$
$$\Gamma x_9 = \{x_1, x_2, x_3, x_4, x_7\};$$
$$\Gamma x_{10} = \{x_1, x_2, x_3, x_4, x_5, x_7, x_9\}.$$

Pe baza ansamblului de indicatori X adaptat şi aplicaţiei Γx de mai sus, construim graful $G = (X, \Gamma)$ din fig. 17.

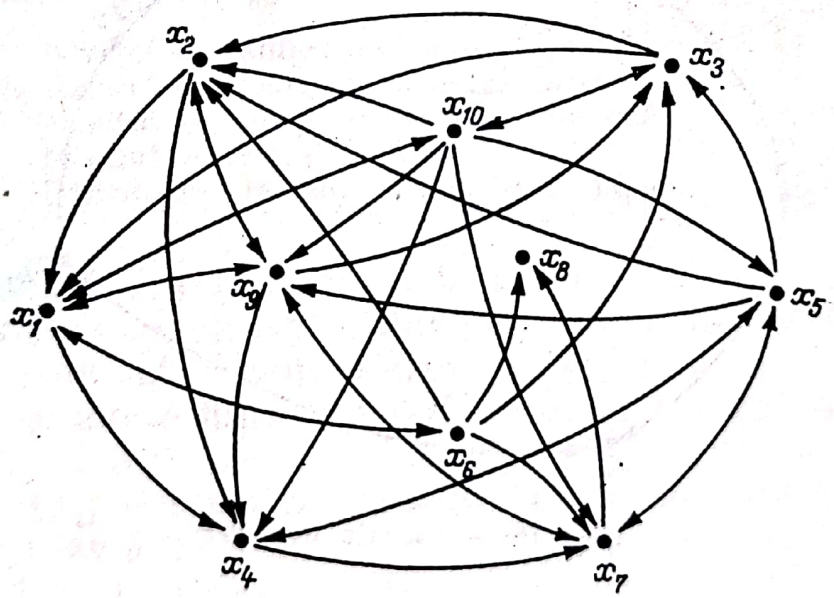


Fig. 17. — Graf orientat

Graful $G = (X, \Gamma)$ din fig. 17 este alcătuit din mai multe lanţuri de arce adiacente, care se sprijină pe nodurile grafului. Aceste lanţuri sînt parcurse în sensul ori-

entării arcelor. Între două puncte ale grafului există sau nu drumuri de lungimi diferite. Cînd drumul revine la punctul de plecare avem un circuit care indică faptul că între indicatorii (variabilele) care intră în reţea există o legătură foarte strînsă. În acest caz, schimbarea fie numai a unui singur indicator, va produce schimbări în toţi indicatorii din reţea.

Cunoaşterea legăturilor foarte strînse existente într-un grafic reţea prezintă un interes deosebit. În vederea lărgirii orizontului cunoaşterii acestor legături într-o anumită structură socială este necesară evidenţierea unor circuite în graficul reţea. De exemplu, în graficul reţea din figura 17, există circuitul (x_1, x_4, x_5, x_3, x_2, x_1). Acest circuit cuprinde următoarele legături foarte strînse între variabile : (x_1, x_4), (x_4, x_5), (x_5, x_3), (x_3, x_2) şi (x_2, x_1) şi este prezentat în figura 18.

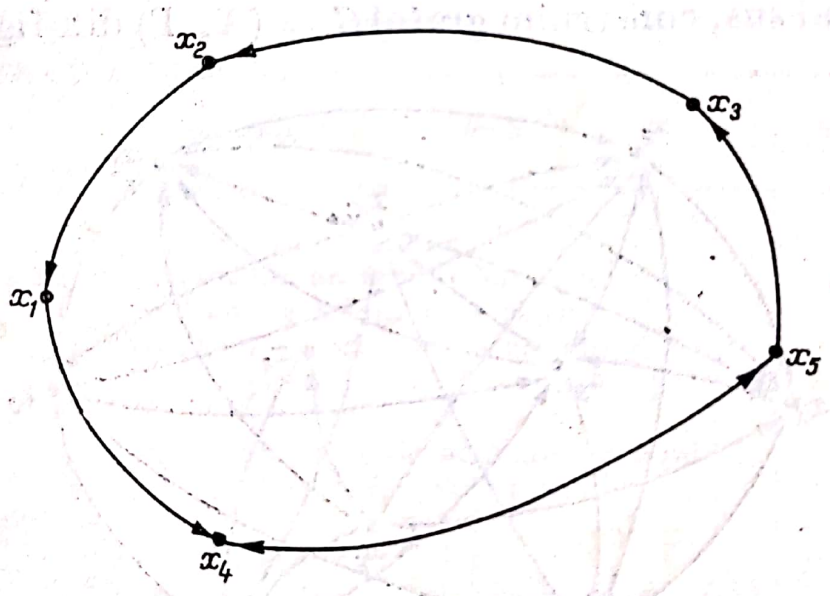


Fig. 18. — Circuit în graf

Circuitul din fig. 18 se interpretează astfel : numărul membrilor de familie (x_1) influenţează numărul persoanelor ocupate în cîmpul muncii (x_4) ; numărul persoanelor ocupate în cîmpul muncii (x_4) determină salariul total al acestora (x_5) ; salariul total al celor aflaţi în cîmpul muncii

(x_5) are o influenţă asupra numărului copiilor de vîrstă preşcolară (x_3); numărul copiilor de vîrstă preşcolară (x_3) influenţează numărul membrilor de familie întreţinuţi (x_2), iar acesta determină numărul membrilor de familie (x_1).

Aceasta nu înseamnă că întotdeauna creşterea numărului membrilor de familie antrenează în mod obligatoriu sporirea numărului celor aflaţi în cîmpul muncii. Totuşi, dacă în familie mai există membri apţi de muncă neîncadraţi în cîmpul muncii, uneori aceştia vor fi nevoiţi să lucreze într-o întreprindere pentru a spori în felul acesta veniturile din salarii ale familiei.

Un alt exemplu de circuit independent de graful din figura 17 se dă în figura 19 pentru sistemul de indicatori din tabelul 3.

Tabelul 3

Nr. crt.	Denumirea indicatorului	Simbol
1	Nivelul de calificare al muncitorului	x_1
2	Nivelul de salarizare al muncitorului	x_2
3	Venitul pe cap de membru al familiei	x_3
4	Nivelul de trai al familiei	x_4
5	Distribuirea în familie a timpului liber	x_5

Circuitul din fig. 19 se interpretează astfel: nivelul de calificare al muncitorului (x_1) are o influenţă asupra nivelului de salarizare al acestuia (x_2); nivelul de salarizare al muncitorului (x_2) influenţează venitul pe cap de membru al familiei (x_3); venitul pe cap de membru al familiei (x_4) determină distribuirea în familie a timpului liber (x_5) iar acesta din urmă influenţează nivelul de calificare al muncitorului (x_1).

De exemplu, cînd creşte nivelul de salarizare al muncitorului, sporesc veniturile familiei şi implicit nivelul ei de trai, iar acesta determină o nouă distribuire a timpului liber în familie.

Scanned with OKEN Scanner

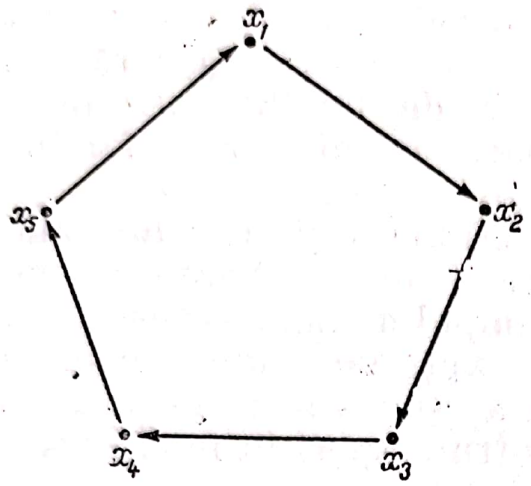


Fig. 19. — Circuit închis circulabil

Prin urmare, într-un astfel de circuit există o legătură foarte strînsă între indicatori.

Într-un graf mai pot exista circuite închise care nu pot fi parcurse în întregime, în sensul orientării arcelor. În graful din fig. 17 se poate evidenţia un astfel de circuit ca în figura 20.

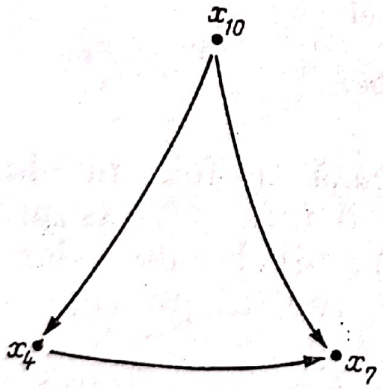


Fig. 20. — Circuit închis necirculabil

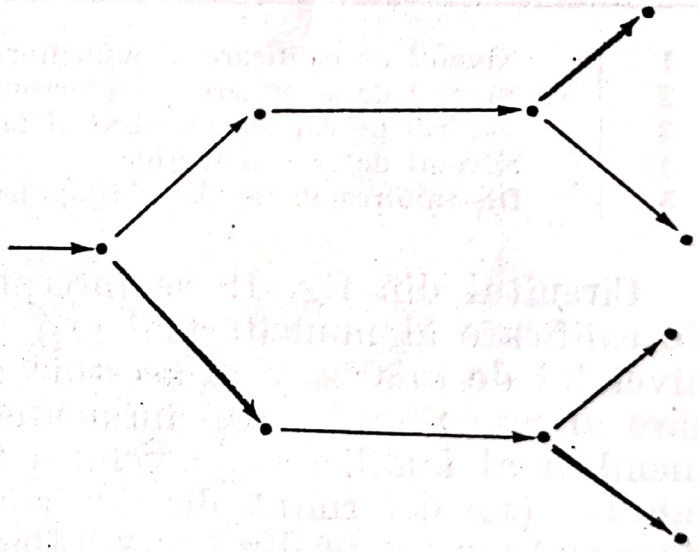


Fig. 21. — Graf arborescent

Circuitul necirculabil din fig. 20 indică faptul că legătura dintre numărul membrilor de familie aflaţi în cîmpul muncii (x_4) şi nivelul de dezvoltare culturală a familiei

(x_7) este influenţată de un indicator comun: costul de trai minim al familiei (x_{10}).

Pe lîngă grafurile cu circuite circulabile sau necirculabile, există grafuri lipsite de astfel de circuite care poartă denumirea de *grafuri arborescente*. Exemplul unui astfel de graf este dat în figura 21.

Graful arborescent din fig. 21 indică faptul că între doi indicatori din reţea există o legătură care este cu atît mai slabă cu cît distanţa măsurată prin numărul arcelor este mai mare.

2.7. MATRICEA ASOCIATĂ UNUI GRAF

Graficele-reţea complexe conţin foarte multe noduri (puncte) care fac imposibilă analiza legăturilor reciproce dintre indicatorii corespunzători acestora, cu ajutorul construcţiilor descrise la 2.6. De aceea, în asemenea cazuri, se apelează la matricea asociată grafului, adică la un model matematic care reproduce exact însuşirile graficului reţea. Fiecărui graf îi corespunde o singură matrice de conexiune, care reflectă toate legăturile cauzale existente între indicatori. De aceea, o astfel de matrice poate fi scrisă astfel:

$$A = (G) = (g_{ij}),$$

în care, g_{ij} reprezintă elementele matricei G.

Elementele g_{ij} ale matricei G indică existenţa sau inexistenţa unor legături în graficul reţea. Dacă X reprezintă ansamblul nodurilor $x_i \in X$ iar Y — ansamblul arcelor $y_{ij} \in Y$, atunci elementele g_{ij} ale matricei $G = (X, Y)$ vor fi:

$$g_{ij} = 1 \text{ dacă } (x_i, x_j) \in Y \text{ şi}$$

$$g_{ij} = 0 \text{ dacă } (x_i, x_j) \not\subset Y.$$

Deoarece G reprezintă un graf, vom nota prin (G) matricea asociată acestuia — un tabel cu dublă intrare, avînd n linii şi n coloane. Un astfel de tabel de legături posibile se prezintă astfel:

Scanned with OKEN Scanner

Variabile care influențează (x_i) \ Variabile influențate (x_i)	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
x_1				1		1			1	1
x_2	1			1		1			1	
x_3	1		1							1
x_4					1		1			
x_5		1	1	1			1		1	
x_6	1	1	1				1	1		
x_7					1			1	1	
x_8										
x_9	1	1	1	1			1			
x_{10}	1	1	1	1	1		1		1	

Tabela acestor legături poate fi scrisă şi sub formă matricială (matricea 2).

$A =$	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
x_1	0	0	0	1	0	1	0	0	1	1
x_2	1	0	0	1	0	1	0	0	1	0
x_3	1	0	1	0	0	0	0	0	0	1
x_4	0	0	0	0	1	0	1	0	0	0
x_5	0	1	1	1	0	0	1	0	1	0
x_6	1	1	1	0	0	0	1	1	0	0
x_7	0	0	0	0	1	0	0	1	1	0
x_8	0	0	0	0	0	0	0	0	0	0
x_9	1	1	1	1	0	0	1	0	0	0
x_{10}	1	1	1	1	1	0	1	0	1	0

Matricea 2. Matricea asociată unui graf (matricea de conexiune)

Scanned with OKEN Scanner

În matricea 2 am notat cu 1 existenţa unei legături cauzale între doi indicatori (x_i, x_j) iar cu 0 lipsa unei asemenea legături. Prin urmare, matricea de conexiune arată indicatorii care influenţează asupra altora.

La construirea matricei de conexiune se au în vedere următoarele principii :

a) se analizează influenţa variabilei x_i asupra variabilei x_j, şi nici de cum legătura reciprocă ;

b) se analizează influenţa variabilei x_i asupra variabilei x_j în ipoteza că toate celelalte variabile au fost considerate la un nivel constant.

Respectarea acestor două principii uşurează procesul de cunoaştere a structurii sociale analizate.

Matricea asociată unui graf înlesneşte studiul unor proprietăţi importante ale grafului. Astfel, cu ajutorul matricei de conexiuni se poate calcula gradul de determinare a unor variabile de către altele, precum urmează :

a) se stabileşte ansamblul arcelor incidente în exteriorul unui nod $x_i \in X$, adică *gradul de emisie* :

$$d^{+}_{(x_i)} = \sum_{j=1}^{n} g_{ij},$$

care reprezintă numărul arcelor din linia matricei de conexiuni corespunzătoare punctului x_i ; de exemplu, în matricea 2, $d^{+}_{(x_1)} = 4$;

b) se stabileşte ansamblul arcelor incidente în interiorul aceluiaş nod $x_i \in X$, adică *gradul de recepţie* :

$$d^{-}_{(x_i)} = \sum_{i=1}^{n} g_{ji},$$

care reprezintă numărul arcelor din coloana matricei corespunzătoare nodului x_i ; de exemplu, în matricea de conexiune din fig. 19, $d^{-}_{(x_1)} = 5$.

c) se calculează diferența între gradul de recepție și gradul de emisie pentru un nod oarecare $x_i \in X$, adică:

$$D_{(x_i)} = d^-_{(x_i)} - d^+_{(x_i)} = \sum_{j=1}^{n} g_{ji} - \sum_{i=1}^{n} g_{ij}.$$

De exemplu, în matricea 2, diferența

$$D_{(x_1)} = d^-_{(x_1)} - d^+_{(x_2)} = 5 - 4 = 1, \text{ iar}$$

$$D_{(x_{10})} = d^-_{(x_{10})} - d^+_{(x_{10})} = 2 - 7 = -5.$$

d) se calculează gradul de determinare al unui indicator de către altul cu ajutorul formulei:

$$I_{(x_i)} = \frac{D_{(x_i)}}{n-1} = \frac{\sum_{j=1}^{n} g_{ji} - \sum_{i=1}^{n} g_{ij}}{n-1},$$

în care:

n — numărul nodurilor existente în graficul rețea.

De exemplu, matricea 2, pentru un număr de puncte $n = 10$ și o diferență între numărul arcelor incidente în exterior și a celor incidente în interior $D_{(x_{10})} = 5$, obținem un grad de determinare a indicatorului x_{10} de către ceilalți indicatori x_j din graficul rețea:

$$I_{(x_{10})} = \frac{D_{(x_{10})}}{n-1} = \frac{-5}{10-1} = -0{,}56$$

Valoarea și semnul negativ al acestui indice reflecta o influență medie a indicatorului x_{10} asupra celorlalți indicatori x_j din sistem, deoarece indicele $I_{(x_i)} = -1$ pune în evidență un factor perfect determinat. Dacă în aceeași

matrice de conexiune luăm în considerare diferența $D_{(x_1)} = 1$, obținem un grad de influență a indicatorilor x_j din graficul rețea asupra indicatorului x_1 :

$$I_{(x_1)} = \frac{D_{(x_1)}}{n-1} = \frac{1}{10-1} = 0{,}11.$$

Valoarea și semnul pozitiv al acestui indice evidențiază faptul că indicatorul x_1 este determinat relativ slab de ceilalți indicatori x_j din graficul rețea studiat.

În virtutea celor arătate mai înainte, în studiul structurilor sociale, semnul și valoarea indicelui $I_{(x_i)}$ caracterizează gradul de determinare al unui indicator x_i de către ceilalți indicatori x_j, precum urmează :

a). indicii $I_{(x_i)} = -0{,}71, \ldots, -1{,}00$ semnalează prezența unei influențe puternice a indicatorului x_i asupra celorlalți indicatori x_j ; indicatorul x_i devine, în asemenea condiții, determinant în structura socială studiată. De exemplu, în matricea 2 nu întîlnim o astfel de relație cauzală ;

b) indicii $I_{(x_i)} = 0{,}31, \ldots, -0{,}70$ indică prezența unei influențe medii a indicatorului x_i asupra celorlalți indicatori x_j ; indicatorul x_i este, în asemenea condiții, determinant într-o măsură mai redusă decît în cazul precedent. De exemplu, în matricea 2, indicii $I_{(x_{10})} = -0{,}56$ și $I_{(x_6)} = -0{,}33$ semnalează faptul că indicatorii x_{10} și x_6 exercită o influență medie asupra celorlalți indicatori din structura socială studiată ;

c) indicii $I_{(x_i)} = 0, \ldots, -0{,}30$ evidențiază o influență foarte slabă a indicatorului x_i asupra celorlalți indicatori x_j din graficul rețea ; indicatorul x_i este determinant într-o măsură redusă. De exemplu, în matricea 2, indicii $I_{(x_2)} = I_{(x_9)} = 0$ semnalează faptul că acești indicatori sînt în egală măsură determinanți și independenți ; indicele $I_{(x_5)} = 0{,}22$ evidențiază faptul că indicatorul x_5 exercită o influență slabă asupra celorlalți indicatori din structura socială analizată ;

Scanned with OKEN Scanner

d) indicii $I_{(x_i)} = 0, \ldots, +0{,}30$ pun în evidenţă existenţa unei determinări slabe a indicatorului x_i de către ceilalţi indicatori x_j din sistem. De exemplu în matricea 2, indicii $I_{(x_i)} = +0{,}11$, $I_{(x_3)} = I_{(x_8)} = +0{,}22$ evidenţiază faptul că indicatorii x_1, x_3, x_7 şi x_8 sînt slab determinaţi de ceilalţi indicatori în structura socială analizată ;

e) indicii $I_{(x_i)} = +0{,}31, \ldots +0{,}70$ semnalează existenţa unei determinări medii a indicatorului x_i de către ceilalţi indicatori x_j din sistem. De exemplu, în matricea 2, indicii $I_{(x_4)} = +0{,}33$ indică faptul că indicatorul x_4 este determinat la un nivel mediu de către ceilalţi indicatori ;

f) indicii $I_{(x_i)} = +0{,}71, \ldots, +1{,}00$ evidenţiază o determinare foarte puternică a indicatorului x_i de către ceilalţi indicatori x_j din sistem. În matricea 2 nu există o astfel de legătură.

Cunoscîndu-se factorii determinanţi, principali ai unei structuri sociale putem stabili relaţiile cauzale esenţiale şi neesenţiale dintre variabilele ce caracterizează această structură. Dacă admitem ipoteza că o cauză determinantă într-o anumită structură socială va produce un efect determinant, atunci relaţiile cauzale existente între o asemenea variabilă şi celelalte variabile din sistem vor fi esenţiale într-o astfel de structură, iar restul vor fi neesenţiale [3]. De exemplu, pentru structura din figura 17, costul de trai minim al familiei (x_{10}) este un indicator determinant pentru structura luată în studiu, exercitînd un efect preponderent asupra ansamblului indicatorilor : numărul membrilor din familie (x_1), numărul membrilor întreţinuţi din familie (x_2), numărul copiilor de vîrstă preşcolară (x_3), numărul persoanelor aflate în cîmpul muncii (x_4), salariul total al persoanelor aflate în cîmpul muncii (x_5), nivelul de dezvoltare culturală al membrilor familiei (x_7) şi veniturile familiei, cu excepţia salariilor (x_9). Relaţiile cauzale care

[3] Într-o structură socială se pot întîlni mai multe cauze determinante care să genereze, fiecare în parte, relaţii cauzale esenţiale.

se stabilesc, în acest caz, între acești indicatori și costul de trai minim al familiei (x_{10}) pot fi considerate esențiale pentru structura socială luată în studiu.

Cînd se analizează un factor determinant dintr-o structură socială oarecare, relațiile cauzale esențiale sînt acelea care determină variabila examinată. De exemplu, dacă studiem factorii care condiționează numărul persoanelor din familie aflate în cîmpul muncii (x_4), observăm că în fig. 17 acest indicator este determinat de numărul membrilor din familie (x_1), numărul membrilor întreținuți din familie (x_2), salariul total al persoanelor aflate în cîmpul muncii (x_5), veniturile familiei, cu excepția salariilor, (x_9) și costul de trai minim al familiei (x_{10}). Desigur că, în acest caz particular, relațiile cauzale pot fi considerate esențiale în analiza factorilor care determină numărul persoanelor aflate în cîmpul muncii (x_4).

Acest mod de deducere a relațiilor cauzale esențiale și neesențiale dintre variabile, trebuie privit însă cu multe rezerve. El nu poate fi acceptat decît în faza de emitere a ipotezelor asupra acestor relații cauzale, verificarea previziunilor făcute fiind obligatorie în toate cazurile, prin cuantificarea indicatorilor, măsurarea corelațiilor și a inferențelor cauzale, cu ajutorul metodelor ce se vor prezenta în capitolele 3 și 4.

2.8. METODA DE CERCETARE A DEPĂRTĂRILOR ÎNTR-UN GRAF

Lungimea unui drum δ într-un graf, pe care-l notăm cu $e(\delta)$, este, prin definiție, numărul de arce ce formează drumul examinat. Dacă între două noduri x_i și x_j există unul sau mai multe drumuri $\delta_m(x_i, x_j)$, depărtarea dintre x_i și x_j este, prin definiție :

$$e(x_i, x_j) = \min \{e[\delta_m(x_i, x_j)]\},$$

Scanned with OKEN Scanner

reprezentînd cel mai scurt drum între x_i şi x_j. Prin urmare, depărtarea dintre x_i şi x_j va fi $e(x_i, x_j)=1,2,\ldots,n$ arce care formează drumul analizat ; iar dacă nu există niciun drum, $e\ (x_i,\ x_j) = \infty$.

Prezenţa unui drum între punctele x_i şi x_j poate fi pusă în evidenţă de arcele parcurse în sens unic, adică în sensul indicat prin săgeţi. În caz contrar, cînd nu este satisfăcută această condiţie, nu există nici un drum.

Pentru a fi posibilă urmărirea uşoară a depărtărilor între toate nodurile grafului $G = (X, \Gamma)$, folosim un algoritm de calcul. În acest scop, se utilizează o metodă matricială, care este mai riguroasă. De exemplu, pentru graful din figura 17 se poate construi matricea 3.

$E=$	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
x_1	0	2	2	1	2	1	2	2	1	1
x_2	1	0	2	1	3	1	2	2	1	2
x_3	1	1	0	2	2	2	2	3	2	1
x_4	3	2	2	0	1	3	1	2	2	3
x_5	2	1	1	1	0	2	1	2	1	2
x_6	1	1	1	2	2	0	1	1	2	2
x_7	2	2	2	2	1	3	0	1	1	3
x_8	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞
x_9	1	1	1	1	2	2	1	2	0	2
x_{10}	1	1	1	1	1	2	1	2	1	0

Matricea 3.—Matricea depărtărilor (ecarturilor)

Cu ajutorul matricei 3 putem stabili toate depărtările posibile dintre punctele grafului, luate pe perechi de puncte. De exemplu, depărtarea dintre x_1 şi x_3 este $e(x_1, x_3)=2$, între x_4 şi x_1 este $e(x_4, x_1)=3$, iar între x_{10} şi x_5 este $e\ (x_{10}, x_1)=1$.

Matricea depărtărilor E permite să se determine :

Scanned with OKEN Scanner

a) modul de transmitere a informaţiilor de la o variabilă la alta (direct, indirect);

b) gradul de conexiune al grafului (conexitatea grafului) care reflectă densitatea legăturilor existente în structura socială analizată.

Prin studiul modului de transmitere al informaţiilor de la o variabilă la alta putem sesiza orientativ, aproximativ, atît legăturile cele mai puternice cît şi cele mai slabe. Dacă dorim să cunoaştem cu exactitate trăsăturile legăturilor existente într-un graf, apelăm la serviciile analizei corelaţionale şi ale analizei de dependenţă, care permit măsurarea relaţiilor statistice dintre indicatorii cuantificaţi în prealabil. Examinînd matricea depărtărilor E din fig. 22, observăm că depărtarea unui nod x_j din graf faţă de altul x_i poate fi $e\ (x_i,\ x_j) = 0, 1, 2, 3, \ldots, \infty$. Fiecare dintre aceste valori ale depărtării $e(x_i, x_j)$ determină un anumit mod de transmitere a informaţiilor de la o variabilă x_i la alta x_j:

a) depărtarea $e\ (x_i, x_j) = 0$ indică gruparea nodurilor x_i şi x_j într-unul singur; acest caz l-am întîlnit în construirea grafului redus din figura 10, obţinut dintr-un graf obişnuit în care unele noduri s-au grupat într-unul singur;

b) depărtarea $e\ (x_i, x_j)=1$ indică faptul că indicatorul x_i influenţează direct indicatorul x_j. De exemplu, în graful din figura 17 depărtarea $e\ (x_3, x_2)=1$ evidenţiază faptul că numărul copiilor de vîrstă preşcolară (x_3) are o influenţă directă asupra numărului membrilor de familie întreţinuţi (x_2);

c) depărtarea $e(x_i, x_j) = 2$ indică faptul că indicatorul x_i influenţează indirect indicatorul x_j. De exemplu, în graful din figura 17, depărtarea $e(x_3, x_4)=2$ evidenţiază faptul că numărul copiilor de vîrstă preşcolară (x_3) influenţează indirect numărul celor aflaţi în cîmpul muncii (x_4), deoarece între aceşti doi indicatori se interpune indicatorul ce caracterizează costul de trai minim al familiei (x_{10}), care acţionează direct asupra numărului celor aflaţi în cîmpul muncii (x_4);

Scanned with OKEN Scanner

d) depărtarea $e(x_i, x_j) = 3, 4, 5, \ldots, n$ indică legături indirecte foarte slabe sau aproape inexistente între variabilele considerate. De exemplu, în graful din figura 17, depărtarea $e(x_1, x_8) = 3$ evidenţiază faptul că între numărul persoanelor din familie (x_1) şi sporul la învăţătură al copiilor (x_8) există o conexiune slabă;

e) depărtarea $e(x_i, x_j) = \infty$ indică lipsa unei legături directe sau indirecte între indicatorii consideraţi. De exemplu, în graful din figura 17, depărtarea $e(x_8, x_6) = \infty$ indică lipsa unui drum şi prin urmare, inexistenţa influenţei sporului la învăţătură al copiilor (x_8) asupra măririi suprafeţei locuibile ocupate de familie (x_6).

Potrivit celor arătate mai înainte, cu cît drumul dintre două noduri x_i şi x_j din graficul reţea are o lungime mai mare (măsurată prin numărul arcelor pe care le conţine), cu atît legătura dintre indicatorii corespunzători acestor noduri este mai slabă. În caz contrar, cu cît drumul dintre două noduri în graficul reţea este mai scurt, cu atît legătura dintre indicatorii corespunzători acestor noduri este mai puternică.

După ce am analizat modul de transmitere al informaţiilor între perechile de indicatori (x_i, x_j) cu ajutorul matricei depărtărilor, trebuie să subliniem gradul de conexiune al grafului (conexitatea grafului).

Noţiunea de conexitate este deosebit de importantă în studiul relaţiilor cauzale dintre variabilele unei structuri sociale. Astfel, prin determinarea conexiunii unei structuri sociale se estimează densitatea legăturilor cauzale existente în structura socială studiată. Această densitate a legăturilor existente într-o astfel de structură se poate aprecia cu ajutorul *matricei depărtărilor* [7].

În vederea stabilirii gradului de conexiune al grafului se poate utiliza unul din următoarele procedee:

a) determinarea indicelui de conexiune (conexitate) al grafului cu ajutorul matricei de conexiune;

b) interpretarea intuitivă a matricei depărtărilor.

În cazul primului procedeu, indicele de conexiune al grafului se defineşte ca raport între numărul legăturilor

efective existente între variabile (N_e) şi numărul total posibil al relaţiilor cauzale dintre variabile (N_p), adică prin :

$$I_c = \frac{N_e}{N_p} = \frac{N_e}{n(n-1)},$$

în care :

n — numărul indicatorilor (variabilelor) structurii sociale.

Acest indice se calculează pe baza elementelor matricei depărtărilor, prin însumarea tuturor legăturilor directe existente între indicatori (N_e).

În funcţie de valoarea indicelui de conexiune I_c al graficului reţea se apreciază gradul de conexiune al structurii sociale analizate ;

— indicii $I_c = 0{,}67, \ldots, 1{,}00$ evidenţiază o structură tare conexă ;

— indicii $I_c = 0{,}34, \ldots, 0{,}67$ reflectă o structură socială conexă ;

— indicii $I_c = 0, \ldots, 0{,}33$ evidenţiază o structură socială slab conexă.

De exemplu, pentru matricea 3 indicele de conexiune al grafului este :

$$I_c = \frac{N_e}{n(n-1)} = \frac{38}{I_c(10-1)} = \frac{38}{90} = 0{,}42.$$

Deoarece valoarea acestui indice se încadrează între limitele lui $I_c = 0{,}34, \ldots, 0{,}67$, graful este considerat conex.

În cel de al doilea caz, cînd se foloseşte interpretarea intuitivă a matricei depărtărilor, aprecierea gradului de conexiune a structurii sociale se face în funcţie de frecvenţa *depărtărilor* $e(x_i, x_j) = \infty$ (lipsite de drum) :

a) lipsa sau frecvenţa foarte redusă a relaţiilor cauzale cu depărtarea $e(x_i, x_j) = \infty$ evidenţiază o structură socială tare conexă ;

Scanned with OKEN Scanner

b) frecvenţa redusă a relaţiilor cauzale avînd depărtarea $e(x_i, x_j) = \infty$ reflectă o structură socială conexă;

c) frecvenţa mare sau foarte mare a relaţiilor cauzale avînd depărtarea $e(x_i, x_j) = \infty$ evidenţiază o structură socială slab conexă.

Acest procedeu conduce la rezultate mai puţin precise decît primul, însă permite aprecierea operativă a conexităţii unei structuri sociale. De exemplu, în matricea de conexiune din fig. 20, se remarcă o frecvenţă foarte redusă a legăturilor avînd depărtarea $e(x_i, x_j) = \infty$, ceea ce evidenţiază o structură socială conexă.

2.9. UN MODEL INDUCTIV APLICAT ÎN STUDIUL PROCESULUI DE MIGRAŢIE A POPULAŢIEI

Să admitem că procesul social de migraţie a populaţiei de la sat la oraş, ca urmare a procesului de industrializare, este alcătuit din mai multe procese particulare, caracterizate de indicatorii din tabelul 4.

Tabelul 4

Procesul social particular		Indicatorii procesului social particular		
Nr. crt.	Denumirea	Nr. crt.	Denumirea	Simbol
1	Procesul economic	1	Creşterea gradului de industrializare urbană	x_1
		2	Nivelul de trai familial	x_2
		3	Ataşament faţă de locuinţă	x_3
2	Procesul geografic	4	Distanţa de migraţie	x_4

Procesul social particular		Indicatorii procesului social particular		
Nr. crt.	Denumirea	Nr. crt.	Denumirea	Simbol
3	Procesul biologic, demografic	5	Frecvența migrației pe unitatea de suprafață	x_5
		6	Structura pe vîrste a populației care migrează	x_6
4	Procesul cultural	7	Nivelul de cultură al populației care migrează	x_7
		8	Gradul de eterogenitate socială și etnică	x_8
5	Procesul psihic	9	Gradul de adaptabilitate la locul de muncă	x_9
6	Procesul juridic	10	Ponderea abaterilor disciplinare	x_{10}
7	Procesul istoric	11	Dinamica migrației	x_{11}
8	Procesul politic	12	Creșterea greutății specifice a muncitorilor în mediul urban	x_{12}

Apreciem că între acești indicatori există următoarele relații de corespondență:

$$\Gamma x_1 = \{x_2, x_4, x_5, x_7, x_9, x_{11}, x_{12}\};$$

$$\Gamma x_2 = \{x_5, x_7, x_9, x_{11}\};$$

$$\Gamma x_3 = \{x_4, x_5, x_{11}\};$$

$$\Gamma x_4 = \{x_5\};$$

$$\Gamma x_5 = \{x_{10}, x_{11}, x_{12}\};$$

$$\Gamma x_6 = \{x_3,\ x_4,\ x_5,\ x_9,\ x_{10}\};$$

$$\Gamma x_7 = \{x_9\};$$

$$\Gamma x_8 = \{x_5\};$$

$$\Gamma x_9 = \{x_5,\ x_{11}\};$$

$$\Gamma x_{10} = \{0\};$$

$$\Gamma x_{11} = \{x_{12}\};$$

$$\Gamma x_{12} = \{0\}.$$

Pe baza grupei de indicatori și a relațiilor de corespondență menționate, construim graful din figura 22.

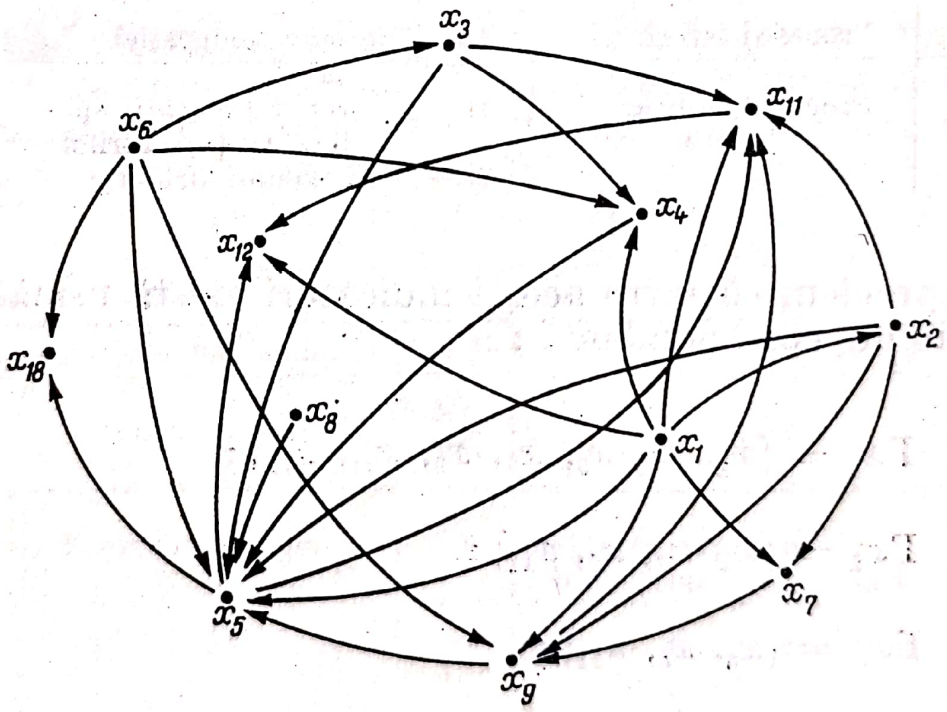


Fig. 22. — Graful legăturilor

Grafului din figura 22 i se poate asocia matricea 4.

$A =$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
x_1	0	1	0	1	1	0	1	0	1	0	1	1
x_2	0	0	0	0	1	0	1	0	1	0	1	0
x_3	0	0	0	1	1	0	0	0	0	0	1	0
x_4	0	0	0	0	1	0	0	0	0	0	0	0
x_5	0	0	0	0	0	0	0	0	0	1	1	1
x_6	0	0	1	1	1	0	0	0	1	1	0	0
x_7	0	0	0	0	0	0	0	0	1	0	0	0
x_8	0	0	0	0	1	0	0	0	0	0	0	0
x_9	0	0	0	0	1	0	0	0	0	0	1	0
x_{10}	0	0	0	0	0	0	0	0	0	0	0	0
x_{11}	0	0	0	0	0	0	0	0	0	0	0	1
x_{12}	0	0	0	0	0	0	0	0	0	0	0	0

Matricea 4. — Matrice de conexiune

Pe baza diferenţelor între elementele liniilor şi ale coloanelor matricei 3, calculăm gradul de influenţă al fiecărui indicator asupra celorlalţi (tabelul 5).

Examinînd indicii din ultima coloană a acestui tabel, observăm că indicele $I_{(x_1)} = -0{,}64$ are cea mai mare valoare negativă, aşadar creşterea gradului de industrializare urbană (x_1) constituie factorul determinant al migraţiei populaţiei de la sat la oraş, deoarece el exercită cea mai mare influenţă asupra celorlalţi indicatori din graficul reţea. Întrucît acest indice este cuprins în intervalul lui $I_{(x_i)} = -0{,}31 \ldots -0{,}70$, evidenţiază faptul că indicatorul x_1 exercită o influenţă medie asupra celorlalţi indicatori din graficul reţea (x_4, x_5, x_7, x_9, x_{11} şi x_{12}). O influenţă similară o exercită asupra structurii sociale analizate şi indicatorul structurii pe vîrstă a populaţiei (x_6), avînd un grad de dependenţă : $I_{(x_8)} = -0{,}45$.

Scanned with OKEN Scanner

Tabelul 5

Nr. crt.	Denumirea indicatorului	Simbolul	Numărul arcelor incidente		col.3 — col. 4.	Gradul de influență
			în exterior	în interior		col. 5
						$n-1$
0	1	2	3	4	5	6
1	Gradul de industrializare urbană	x_1	7	0	7	−0,64
2	Nivelul de trai familial	x_2	4	1	3	−0,27
3	Atașament față de locuință	x_3	3	1	2	−0,18
4	Distanța de migrație	x_4	1	3	−2	+0,18
5	Frecvența migrației pe unitatea de suprafață	x_5	3	7	−4	+0,36
6	Structura pe vîrste a populației care migrează	x_6	5	0	5	−0,45
7	Nivelul de cultură al persoanelor care migrează	x_7	1	2	−1	+0,09
8	Gradul de eterogenitate socială și etnică	x_8	1	0	1	−0,09
9	Gradul de adaptabilitate la locul de muncă	x_9	2	4	−2	+0,18
10	Ponderea abaterilor disciplinare	x_{10}	0	2	−2	+0,18
11	Dinamica migrației	x_{11}	1	5	−4	+0,36
12	Greutatea specifică a muncitorilor din mediul urban	x_{12}	0	3	−3	−0,27

Pe lîngă cei doi indicatori menționați, o influență slabă asupra celorlalți indicatori o exercită : nivelul de

trai familial (x_2), atașamentul față de locuință (x_3) și gradul de eterogenitate socială și etnică (x_8).

Factorii determinați într-o măsură diferită în structura socială analizată, reflectați de valorile pozitive ale indicelui ce evidențiază gradul de dependență, sînt următorii: frecvența migrației pe unitatea de suprafață (x_5), dinamica migrației (x_{11}), creșterea greutății specifice a muncitorilor în mediul urban (x_{12}), gradul de adaptabilitate la locul de muncă (x_9), ponderea abaterilor disciplinare (x_{10}) și distanța de migrație (x_5).

Tot cu ajutorul matricei 4 putem determina atît legăturile cauzale esențiale, cît și cele neesențiale. În ipoteza în care o cauză determinantă produce un efect esențial, rezultă că indicatorul determinant ce reflectă creșterea gradului de industrializare urbană (x_1) exercită un efect important asupra structurii sociale analizate, iar influențele lui asupra altor indicatori de care este legat, devin esențiale în această structură:

$$\Gamma x_1 = \{x_2, x_4, x_5, x_7, x_9, x_{11}, x_{12}\}.$$

Această relație evidențiază influența determinantă a creșterii gradului de industrializare urbană (x_1) asupra nivelului de trai familial (x_2), asupra distanței de migrație (x_4), asupra frecvenței migrației pe unitatea de suprafață (x_5), asupra nivelului de cultură al persoanei care migrează (x_7), asupra gradului de adaptabilitate la locul de muncă (x_9), asupra dinamicii migrației populației (x_{11}) și asupra creșterii greutății specifice a muncitorilor din mediul urban (x_{12}). În afara acestor legături, prezintă interes și următoarele relații de corespondență:

$$\Gamma x_6 = \{x_3, x_4, x_5, x_9, x_{10}\};$$

$$\Gamma x_2 = \{x_5, x_7, x_9, x_{11}\};$$

$$\Gamma x_3 = \{x_4, x_5, x_{11}\};$$

$$\Gamma x_8 = \{x_5\}.$$

Scanned with OKEN Scanner

De exemplu, legăturile existente între structura pe vîrste a populaţiei ce migrează (x_6) şi ataşamentul faţă de locuinţă (x_3); nivelul de trai familial (x_2) şi frecvenţa migraţiei pe unitatea de suprafaţă (x_5); ataşamentul faţă de locuinţă (x_3) şi frecvenţa migraţiei pe unitatea de suprafaţă (x_5) sau gradul de eterogenitate socială, etnică şi frecvenţa migraţiei (x_5) etc. pot fi considerate legături esenţiale.

În afara legăturilor esenţiale, ne interesează uneori şi legăturile neesenţiale, care sînt evidenţiate de influenţele indicatorilor determinaţi asupra celorlalţi indicatori din structura socială analizată:

$$\Gamma x_7 = \{x_9\};$$

$$\Gamma x_4 = \{x_5\};$$

$$\Gamma x_9 = \{x_5, x_{11}\};$$

$$\Gamma x_5 = \{x_{10}, x_{11}, x_{12}\};$$

$$\Gamma x_{11} = \{x_{12}\}.$$

De exemplu, influenţa migraţiei populaţiei (x_5) asupra ponderii abaterilor disciplinare (x_{10}) reprezintă o legătură cauzală neesenţială.

După ce am stabilit factorul determinant al structurii social analizate, calculăm depărtările existente în graful din figura 22, care oferă posibilitatea evidenţierii unor particularităţi ale acestei structuri deosebit de complexe. Pentru aceasta folosim matricea 5.

Elementele matricii 5 indică valorile depărtărilor existente în graful din figura 22 pe baza cărora se poate aprecia modul de transmitere al informaţiilor de la o variabilă la alta. Astfel, în funcţie de mărimea depărtării pot fi puse în evidenţă următoarele tipuri de legături:

Scanned with OKEN Scanner

E	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
x_1	0	1	∞	1	1	∞	1	∞	1	2	1	1
x_2	∞	0	∞	∞	1	∞	1	∞	1	2	1	2
x_3	∞	∞	0	1	1	∞	∞	∞	∞	2	1	2
x_4	∞	∞	∞	0	1	∞	∞	∞	∞	2	2	2
x_5	∞	∞	∞	∞	0	∞	∞	∞	∞	1	1	1
x_6	∞	∞	1	1	1	0	∞	∞	1	1	2	2
x_7	∞	∞	∞	∞	2	∞	0	∞	1	3	2	3
x_8	∞	∞	∞	∞	1	∞	∞	0	∞	2	2	2
x_9	∞	∞	∞	∞	1	∞	∞	∞	0	2	1	2
x_{10}	∞	∞	∞	∞	∞	∞	∞	∞	∞	0	∞	∞
x_{11}	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	0	1
x_{12}	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	0

Matricea 5. — Matricea depărtărilor

a) Relaţii directe (x_1, x_2), (x_1, x_4), (x_1, x_3), (x_1, x_7), (x_1, x_9), (x_1, x_{11}), (x_2, x_7), (x_2, x_9), (x_2, x_{11}), (x_3, x_4), (x_3, x_5) ş.a.m.d. Aceste legături se pot interpreta astfel : creşterea gradului de industrializare urbană (x_1) transmite în mod direct informaţii indicatorilor privind frecvenţa migraţiei pe unitatea de suprafaţă (x_5), nivelul de cultură al persoanei care migrează (x_7), gradul de adaptabilitate la locul de muncă (x_9), dinamica migraţiei (x_{11}) şi creşterea greutăţii specifice a muncitorilor în mediul urban (x_{12}); nivelul de trai familial (x_2) transmite în mod direct informaţii indicatorilor privind frecvenţa migraţiei pe unitatea de suprafaţă (x_5), nivelul de cultură al persoanei care migrează (x_7), gradul de adaptabilitate la locul de muncă (x_9) ş.a.m.d.

În graficul reţea din fig. 22 între aceşti indicatori există o depărtare $e(x_i, x_j) = 1$, deoarece influenţa unui indicator x_i asupra altuia x_j este marcată cu un singur arc avînd

săgeata îndreptată în sensul de parcurgere al drumului; această valoare a depărtării (ecartului) indică o legătură directă între indicatorii analizaţi.

b) Relaţii indirecte de ordinul 2 între următorii indicatori:

$$(x_1, x_{10}), (x_2, x_{10}), (x_2, x_{12}), (x_3, x_{10}),$$
$$(x_3, x_{12}), (x_4, x_{10}), (x_4, x_{11}) \text{ ş.a.m.d.}$$

Aceste legături pot fi interpretate astfel: creşterea gradului de industrializare urbană (x_1) transmite în mod indirect informaţii indicatorilor privind: ponderea abaterilor disciplinare (x_{10}), nivelul de trai familial (x_2), abaterile disciplinare (x_{10}) şi creşterea greutăţii specifice a muncitorilor în mediul urban (x_{12}); ataşamentul faţă de locuinţă (x_3) influenţează indirect ponderea abaterilor disciplinare şi creşterea greutăţii specifice a muncitorilor în mediul urban (x_{12}); distanţa de migraţie (x_4) influenţează indirect ponderea abaterilor disciplinare (x_{10}), dinamica migraţiei (x_{11}), creşterea greutăţii specifice a muncitorilor din mediul urban (x_{12}) ş.a.m.d. În graficul reţea din figura 22 între aceşti indicatori există o depărtare $e(x_i, x_j) = 2$, deoarece influenţa unui indicator x_i asupra altuia x_j este marcată cu două arce adiacente ce pot fi parcurse în sensul unic indicat de săgeţi; această valoare a depărtării evidenţiază o legătură indirectă între indicatorii respectivi.

c) Relaţii indirecte de ordin mai mare decît 2 între următorii indicatori: (x_7, x_{10}) şi (x_7, x_{12}). Aceste legături pot fi interpretate astfel: nivelul de cultură al persoanei care migrează (x_7) transmite informaţii foarte slabe indicatorilor privind ponderea abaterilor disciplinare (x_{10}) şi creşterea greutăţii specifice a muncitorilor din mediul urban (x_{12}).

În graficul reţea din figura 22 între aceşti indicatori există o depărtare $e(x_i, x_j) = 3$, deoarece influenţa unui indicator x_i asupra altuia x_j este marcată de trei arce adiacente ce pot fi parcurse în sensul unic indicat de săgeţi; această valoare a depărtării evidenţiază o legătură foarte slabă între indicatorii analizaţi.

Scanned with OKEN Scanner

d) Inexistența unor relații directe sau indirecte între diverși indicatori (x_i, x_j) este evidențiată de valoarea depărtării $e(x_i, x_j) = \infty$, care se poate citi direct în matricea 4. De exemplu, ponderea abaterilor disciplinare (x_{10}) nu influențează în niciun fel creșterea gradului de industrializare urbană (x_1), nivelul de trai familial (x_2), atașamentul față de muncă (x_3) ș.a.m.d.

După ce s-a stabilit modul de transmitere a informațiilor în structura socială analizată, tot cu ajutorul matricei de conexiune din fig. 22 putem determina gradul de conexiune al grafului, care reflectă coeziunea structurii sociale analizate. În acest scop se calculează indicele de conexiune al grafului (I_c) pe baza elementelor matricei depărtărilor, prin însumarea tuturor legăturilor existente între indicatori (N_e), după cum urmează :

$$I_c = \frac{N_e}{n\,(n-1)} = \frac{28}{12\,(12-1)} = 0{,}21.$$

Deoarece indicele I_c se înscrie în intervalul 0 ... 0,33, el reflectă o structură socială slab conexă. De altfel, frecvența mare a legăturilor avînd depărtarea $e(x_i, x_j) = \infty$ din matricea 5 evidențiază o structură socială cu o densitate destul de redusă a legăturilor. Acest lucru este confirmat de altfel și de inexistența circuitelor din graful fig. 22.

În analiza structurilor sociale se poate apela la unele concepte din teoria grafelor în faza de emitere a ipotezelor asupra formei și densității relațiilor cauzale.

2.10. UN ALT EXEMPLU : STUDIUL PROCESULUI DE TRANSMITERE A CULTURII DIN MEDIUL URBAN ÎN CEL RURAL

În faza de emitere a ipotezelor de lucru apelăm la metode inductive de analiză a structurilor sociale [8].

Graful relațiilor cauzale poate fi utilizat cu succes la formularea de ipoteze verificabile asupra unui proces

social oarecare. De pildă, dacă analizăm dinamica procesului de transmitere a elementelor de cultură din mediul urban în cel rural, avem în vedere un proces socioeconomic, ale cărui procese particulare (economice, geografice, demografice, social-culturale etc.) pot fi corelate în diferite moduri [9]. De aceea, obiectivul care trebuie atins va consta în verificarea ipotezei că diferite procese economice, geografice, demografice şi social-culturale au un anumit rol într-un proces social complex, cum este cel de transmitere a culturii din mediul urban în cel rural şi că aceste procese particulare sînt corelate între ele. Pentru caracterizarea acestor procese particulare utilizăm un sistem de indicatori delimitaţi cu precizie. De exemplu, pentru caracterizarea procesului de transmitere a culturii din mediul urban în cel rural se pot utiliza indicatorii cu delimitarea ce li se dă în tabelul 6.

Pentru a verifica ipoteza că transmiterea culturii din mediul urban în cel rural este un proces social complex care implică procese particulare (economice, geografice, demografice şi social-culturale) corelate, trebuie să demonstrăm că între indicatorii din tabelul 6 există un sistem de relaţii cauzale sau de raporturi mutuale. În acest scop, în faza de elaborare a ipotezelor verificabile vom construi un graf orientat $G = (X, Y)$, alcătuit din [10]:

a) ansamblul nodurilor (punctelor) X ale cărui elemente sînt indicatorii x_j din tabelul 6:

$$x_j \in X\,;\ i \in N = \{0, 1, 2, \ldots, n\}\,;$$

b) ansamblul arcelor Y ale cărui elemente sînt:

$$y_{ij} \in Y, \qquad i, j \in N, \qquad i \neq j$$

Prin aceste două mulţimi de noduri X şi arce Y vom scrie o lege conform căreia fiecărui element $x_j \in X$ să-i corespundă un ansamblu de arce incidente spre exterior $\Gamma x \in Y$, care poartă denumirea de aplicaţie a ansamblului

Tabelul 6

Procesele particulare ale procesului general				Indicatorii proceselor particulare	
Nr. crt.	Denumire	Nr. crt.	Indicatorul	Simbol	Sfera de cuprindere
0	1	2	3	4	5
1	Economic	1	Condiţiile de urbanizare ale localităţii	x_1	Dotarea localităţii cu echipament tehnico-edilitar şi sanitar Densitatea populaţiei în locuitori/ha Sistemul de deservire social-culturală Gradul de dotare industrială
		2	Nivelul de construcţii	x_2	Tipul materialului de construcţii Tipul materialului pentru acoperiş Tipul duşumelei Numărul de camere
		3	Mecanizarea agriculturii	x_3	Numărul de tractoare/100 ha
		4	Angajarea în industrie	x_4	Numărul de oameni folosiţi în industrie
		5	Nivelul de trai material	x_5	Venituri Alimente Îmbrăcăminte Locuinţe
		6	Dotare cu obiecte tehnice	x_6	Frigider, maşină de spălat, televizor, radio, maşină, motoretă, bicicletă etc.
		7	Folosirea tracţiunii animale ca forţă energetică	x_7	Număr de animale de muncă/100 ha

Scanned with OKEN Scanner

Tabelul 6 (continuare)

0	1	2	3	4	5
2	Geografic	8	Limita izocronală	x_8	Distanța dintre localități (accesibilitate, durata de parcurgere
3	Demografic	9	Structura pe vîrste a populației	x_9	Preponderența anumitor grupe de vîrstă (tineri-bătrîni)
		10	Sporul natural al populației	x_{10}	Sporul natural ($^0/_{00}$)
4	Social-cultural	11	Contacte sociale	x_{11}	Vizite la rude sau vecini, săli de dans, spectacole, restaurante (bodegi)
		12	Condiții afectînd transmiterea culturii din mediul urban în cel rural	x_{12}	Mijloace de comunicație, raportul locuitorilor din mediul urban cu cei din mediul rural, extinderea electrificării, popularitatea mass-media
		13	Condiții de igienă și obiceiuri	x_{13}	Felul în care se transportă apa. Tip instalații sanitare. Tip canalizare. Tip transport gunoi. Consum de săpun
		14	Bogăția de mass-media	x_{14}	Televizor, radio, cinema (intensitate de folosire)
		15	Intensitate de distracție	x_{15}	Frecvența participării la acțiuni distractive, dans, excursii, petreceri
		16	Ponderea intelectualilor în totalul populației	x_{16}	Numărul intelectualilor față de totalul populației
		17	Transmiterea culturii din mediul urban în cel rural	x_{17}	Echipe artistice, turnee, teatru, caravane cinematografice sau filme, conferințe, cărți și difuzarea presei

Scanned with OKEN Scanner

X în Y. Pentru ansamblul indicatorilor $X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}, x_{17}\}$ care caracterizează, după părerea noastră, procesul de transmitere a culturii din mediul urban în cel rural (a se vedea indicatorii din tabelul 6), vom admite ipoteza că aplicația Γ se caracterizează prin :

$\Gamma x_1 = \{x_2, x_4, x_5, x_6, x_7, x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}, x_{17}\}$;

$\Gamma x_2 = \{x_1, x_5, x_{13}\}$;

$\Gamma x_3 = \{x_1, x_2, x_4, x_5, x_7, x_9, x_{11}\}$;

$\Gamma x_4 = \{x_1, x_2, x_3, x_5, x_6, x_7, x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}\}$;

$\Gamma x_5 = \{x_2, x_4, x_6, x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}, x_{17}\}$;

$\Gamma x_6 = \{x_1, x_2, x_4, x_5, x_8, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}, x_{17}\}$;

$\Gamma x_7 = \{x_1, x_3, x_4, x_5, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}, x_{17}\}$;

$\Gamma x_8 = \{x_1, x_2, x_4, x_5, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}, x_{17}\}$;

$\Gamma x_9 = \{x_1, x_4, x_{10}, x_{11}, x_{12}, x_{14}, x_{15}, x_{17}\}$;

$\Gamma x_{10} = \{x_1, x_4, x_5, x_9, x_{15}\}$;

$\Gamma x_{11} = \{x_2, x_{13}, x_{14}, x_{15}, x_{17}\}$;

$\Gamma x_{12} = \{x_1, x_2, x_4, x_5, x_6, x_{11}, x_{13}, x_{14}, x_{15}, x_{16}, x_{17}\}$;

$\Gamma x_{13} = \{x_1, x_2, x_5, x_6, x_9, x_{10}\}$;

$\Gamma x_{14} = \{x_4, x_5, x_6, x_{11}, x_{12}, x_{13}, x_{15}, x_{17}\}$;

$\Gamma x_{15} = \{x_{11}, x_{14}, x_{17}\}$;

Scanned with OKEN Scanner

$\Gamma x_{16} = \{x_1, x_2, x_5, x_6, x_{12}, x_{13}, x_{14}, x_{17}\}$;

$\Gamma x_{17} = \{x_1, x_2, x_5, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}\}$.

Folosind indicatorii X şi aplicaţia Γx se poate construi graful $G = (X, \Gamma)$ din figura 23 în următoarea succesiune [11]:

a) se marchează prin noduri (puncte) indicatorii x_j ai ansamblului X;

b) se trasează arcele (drepte sau curbe) care leagă nodurile (punctele), avînd în vedere aplicaţia Γx specificată mai înainte;

c) se transformă graful obţinut într-unul cît mai compact şi mai clar.

Principalele concepte din teoria grafelor utilizate în studiul structurilor sociale sînt [12]:

a) *drumul* — poate semnala posibilitatea transmiterii informaţiilor între indicatorii din sistem, ca de exemplu $(x_5, x_4, x_3, x_2, x_1)$;

b) *circuitul* — sugerează posibilitatea circulaţiei intense a informaţiilor între indicatorii din sistem, ca de exemplu $(x_3, x_7, x_5, x_4, x_3)$, şi indică relaţii cauzale puternice, esenţiale între variabilele din circuit;

c) *lungimea* unui drum sau unui circuit—arată existenţa unor relaţii cauzale directe sau indirecte, mai tari sau mai slabe, dintre indicatorii x_i, x_j, de exemplu drumul între nodurile x_6 şi x_1 este $(x_6, x_1) = 1$ şi indică o relaţie cauzală directă, iar drumul între nodurile x_3 şi x_6 care este egal cu 2, indică o relaţie indirectă;

d) *distanţa*— reflectă „depărtarea" existentă între variabilele din structura procesului social analizat, adică lungimea drumului cel mai scurt de transmitere a informaţiilor de la o variabilă la alta într-o structură socială concepută *ca o reţea de comunicaţie*. În figura 24, drumul cel mai scurt între x_1 şi x_8 va fi $d(x_1, x_8) = 2$ [13].

e) *graful simetric*—evidenţiază interacţiuni sau influenţe reciproce puternice între variabilele corelate ale structurii

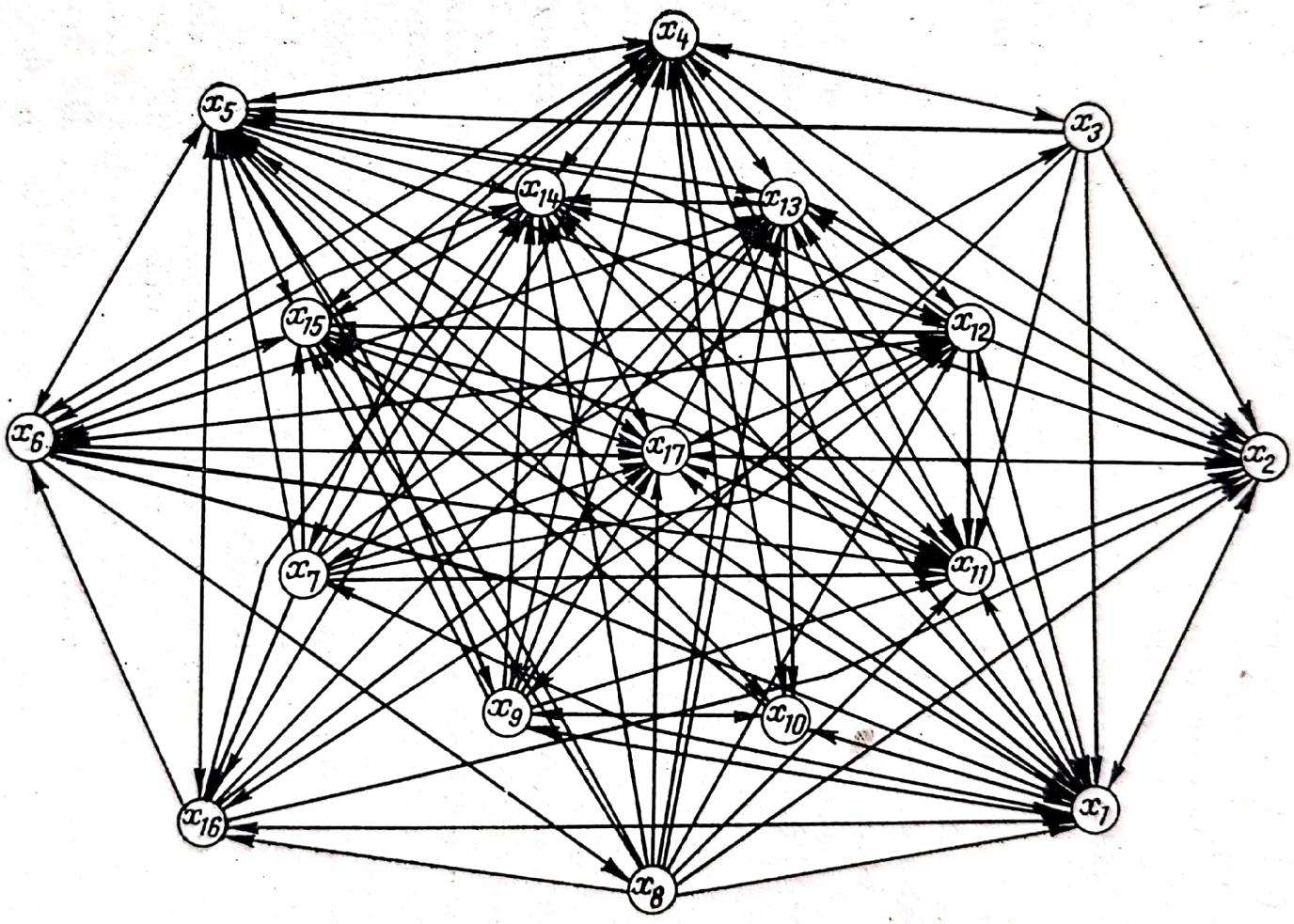


Fig. 23. — Graful legăturilor $G = (X, \Gamma)$

procesului social analizat; graful din figura 23 poate fi considerat parţial simetric, deoarece include legături cauzale directe sau interacţiuni;

f) arborescenţa — graf parţial care nu conţine nici un ciclu (circuit). De exemplu aplicaţia $\Gamma x_3 = \{x_1, x_2, x_4, x_7, x_9, x_{11}\}$ din graful din figura 23, caracterizează un graf parţial avînd un nod central sau de origine (x_3) cu mai multe arce incidente spre exterior, adică spre nodurile x_1, x_2, x_4, x_5, x_7, x_9 şi x_{11}; o arborescenţă de acest fel reflectă capacitatea unei variabile x_j de a transmite informaţii celorlalte variabile x_i din sistem. Cu cît numărul variabilelor x_i va fi mai mare, cu atît cantitatea de informaţie furnizată de variabila x_j va fi mai mare.

Deoarece graful tare conex din fig. 23 evidenţiază un sistem conex de relaţii cauzale sau de raporturi mutuale între variabile putem emite ipoteza că procesul social-economic de transmitere a culturii din mediul urban în cel rural este un proces complex, fiind caracterizat de o serie de procese particulare (economice, geografice, demografice şi social-culturale) care sînt corelate unele cu altele. Această ipoteză va trebui verificată prin efectuarea în continuare a unor observaţii organizate în vederea obţinerii unor rezultate care se pot generaliza cu ocazia finalizării cercetării sociologice.

Interpretarea unor grafuri ale legăturilor cauzale complexe este extrem de dificilă. De aceea, în asemenea cazuri se poate apela cu succes la instrumentul de calcul pe care ni-l furnizează algebra matricială. Astfel grafului orientat din figura 3 i se poate asocia matricea (M) care reproduce exact proprietăţile grafului (matricea 6) [4].

Această matrice este *un tabel cu dublă intrare*, deoarece elementele x_{ij} dispuse pe coloană reflectă *cauza*, iar cele dispuse pe linie ilustrează *efectul* produs în cadrul legăturii existente între două variabile x_i, x_j. De aceea, pentru o variabilă oarecare x_j, elementele x_{ij} dispuse pe linii reflectă efectul produs de variabila x_j asupra celorlalte variabile x_i din sistem, iar cele dispuse pe coloană reflectă efectul produs de variabilele x_i asupra

Scanned with OKEN Scanner

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}
x_1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1
x_2	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0
x_3	1	1	0	1	1	0	1	0	1	0	1	0	0	0	0	0	0
x_4	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	0
x_5	0	1	0	1	0	1	0	0	1	1	1	1	1	1	1	1	1
x_6	1	1	0	1	1	0	0	1	0	0	1	1	1	1	1	0	1
x_7	1	0	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1
x_8	1	1	0	1	1	0	0	0	0	0	1	1	1	1	1	1	1
x_9	1	0	0	1	0	0	0	0	0	1	1	1	0	1	1	0	1
x_{10}	1	0	0	1	1	0	0	0	1	0	0	0	0	0	1	0	0
x_{11}	0	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1
x_{12}	1	1	0	1	1	1	0	0	0	0	1	0	1	1	1	1	1
x_{13}	1	1	0	0	1	1	0	0	1	1	0	0	0	0	0	0	0
x_{14}	0	0	0	1	1	1	0	0	0	0	1	1	1	0	1	0	1
x_{15}	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1
x_{16}	1	1	0	0	1	1	0	0	0	0	0	1	1	1	0	0	1
x_{17}	1	1	0	0	0	1	0	0	0	0	1	1	1	1	1	0	0

Matricea 6. Matricea (M) asociată grafului $G = (X, \Gamma)$

variabilei x_j în cauză. Aşadar, matricea (M) evidenţiază legăturile cauzale directe dintre variabilele x_j, avînd drept corespondent arcele y_{ij} din graful $G = (X, Y)$. Dacă dorim să cunoaştem relaţiile cauzale indirecte vom folosi un algoritm simplu. De pildă, pentru a determina relaţiile cauzale indirecte de distanţă $d(x_i, x_j) = 2, 3, \ldots, n$, vom ridica matricea (M) la puterea $2, 3, \ldots, n$, adică vom folosi matricea $(M)^r$. De exemplu, calculînd matricea $(M)^2$ vom obţine matricea 7, ale cărei elemente ne vor arăta cîte drumuri de lungime 2 există în graful respectiv.

Comparînd elementele din matricele $(M)^2$ şi $(M)^3$ scriem matricea 8, a depărtărilor $E = [\min \{\delta_m e(x_i, x_j)\}]$ [4].

În căsuţa matricei distanţelor E sînt trecute drumurile cele mai scurte, măsurate în număr de arce parcurse, de transmitere a informaţiilor de la o variabilă la alta, într-o structură socială concepută ca o *reţea de comunicaţie* [13].

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}
x_1	10	8	2	7	9	7	1	1	4	4	9	8	10	10	10	4	9
x_2	1	3	0	2	2	3	1	0	3	3	2	2	2	2	2	2	2
x_3	4	4	2	4	4	3	2	0	3	4	5	5	6	6	6	4	5
x_4	9	8	1	9	10	6	2	1	5	4	9	7	9	9	9	4	10
x_5	9	7	1	5	8	6	1	1	3	3	7	6	8	8	8	2	7
x_6	6	8	1	5	7	7	2	0	4	4	8	6	9	8	8	5	7
x_7	6	9	1	5	7	8	3	0	5	4	8	6	8	8	7	4	7
x_8	6	8	1	4	7	8	2	0	4	4	7	6	9	8	7	4	7
x_9	4	5	1	4	5	5	2	0	3	2	6	4	6	6	7	3	5
x_{10}	2	3	1	3	2	3	2	0	3	4	5	4	3	5	4	3	4
x_{11}	3	2	0	1	3	3	0	0	1	1	3	2	3	2	2	0	2
x_{12}	6	8	1	4	7	7	2	1	4	4	7	7	9	8	7	5	7
x_{13}	4	3	0	5	4	2	1	1	3	3	4	4	4	4	5	2	4
x_{14}	5	7	1	3	4	5	1	1	3	3	6	4	6	7	5	3	5
x_{15}	1	2	0	1	1	2	0	0	0	0	2	2	3	2	3	0	2
x_{16}	5	6	0	5	6	5	1	1	3	3	6	5	7	5	5	3	5
x_{17}	4	5	0	4	6	4	1	1	2	2	5	3	6	5	5	2	6

Matricea 7. — Matricea $(M)^2$ a grafului $G = (X, \Gamma)$

Într-o astfel de reţea o variabilă x_j oarecare constituie o sursă de informaţie. Capacitatea unei variabile x_j de a transmite sau de a recepţiona informaţii în sistem va depinde de numărul arcelor y_{ij} incidente spre exterior sau spre

interior. De exemplu, indicatorul ce caracterizează limita izocronală, adică distanţa şi durata de parcurgere a distanţei de la sat la oraş (x_8) va constitui o *variabilă*

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}
x_1	0	1	2	1	1	1	1	2	1	1	1	1	1	1	1	1	1
x_2	1	0	3	2	1	2	2	3	2	2	2	2	1	2	2	2	2
x_3	1	1	0	1	1	2	1	3	1	2	1	2	2	2	2	2	2
x_4	1	1	1	0	1	1	1	2	1	1	1	1	1	1	1	1	2
x_5	2	1	2	1	0	1	2	2	1	1	1	1	1	1	1	1	1
x_6	1	1	2	1	1	0	2	1	2	2	1	1	1	1	1	2	1
x_7	1	2	1	1	1	2	0	3	2	2	1	1	1	1	1	1	1
x_8	1	1	2	1	1	2	2	0	2	2	1	1	1	1	1	1	1
x_9	1	2	2	1	2	2	2	3	0	1	1	1	2	1	1	2	1
x_{10}	1	2	2	1	1	2	2	3	1	0	2	2	2	2	1	2	2
x_{11}	2	1	3	2	2	2	3	3	2	2	0	2	1	1	1	3	1
x_{12}	1	1	2	1	1	1	2	2	2	2	1	0	1	1	1	1	1
x_{13}	1	1	3	2	1	1	2	2	1	1	2	2	0	2	2	2	2
x_{14}	2	2	2	1	1	1	2	2	2	2	1	1	1	0	1	2	1
x_{15}	2	2	3	2	2	2	3	3	3	3	1	2	2	1	0	3	1
x_{16}	1	1	3	2	1	1	2	2	2	2	2	1	1	1	2	0	2
x_{17}	1	1	3	2	2	1	2	2	2	2	1	1	1	1	1	2	0

Matricea 8. — Matricea depărtărilor $E = (e(x_i, x_j))$, unde

$$e(x_i, x_j) = \min \{e[\delta_m(x_i, x_j)]\}$$

determinantă, deoarece numărul arcelor incidente spre exterior este mult mai mare decît al arcelor incidente spre interior; indicatorul ce caracterizează intensitatea distracţiei (x_{15}) va constitui o *variabilă determinată*, deoarece numărul arcelor incidente spre interior este mult mai mare decît al arcelor incidente spre exterior; în fine,

Scanned with OKEN Scanner

Inv. 117649

indicatorul ce caracterizează condiţiile de urbanizare ale localităţii (x_1) va constitui o *variabilă independentă*, adică o variabilă care are capacitatea de a transmite şi de a recepţiona informaţii într-un raport aproximativ egal.

Cînd deducem relaţiile cauzale esenţiale şi neesenţiale din structura socială analizată, cuantificăm în prealabil indicatorii, culegem informaţii de pe teren, prelucrăm aceste informaţii şi calculăm coeficienţii de dependenţă dintre variabile, pe care îi vom analiza în cap. 4. Aceşti coeficienţi reflectă intensitatea relaţiilor cauzale în sensul că dependenţele cele mai puternice dintre variabile vor caracteriza relaţiile cauzale esenţiale iar cele mai slabe — relaţiile cauzale neesenţiale [14], [15].

2.11. ANALIZA RELAŢIILOR DINTRE MEMBRII UNUI GRUP SOCIAL

Investigaţia sociometrică se fundamentează pe ipoteza că între membrii unui grup social există un ansamblu de relaţii [16]. Depistate şi măsurate, aceste relaţii alcătuiesc o structură care poate fi caracterizată geometric şi algebric. Aceste relaţii dau sistemului social structura sa spontană. Ele pot fi atracţii, repulsii, alegeri, indiferenţe, raporturi de dominaţie etc. Reprezentările grafice ale atracţiilor şi repulsiilor iau, într-adevăr, forme tipice: perechi, lanţuri, triunghiuri, cercuri, stele etc. (fig. 23) [17].

În mod curent, *sociograma individuală* se întocmeşte în felul următor: în mijlocul ei se desenează un cerc ce reprezintă un individ oarecare; în jurul acestuia se trec apoi toţi membrii grupului său de care este legat într-un fel sau altul, printr-o serie de linii, fiecare dintre ele avînd semnificaţia ei precisă, dar variind de la cercetător la cercetător [18].

În cercetările sociologice, cercetătorii tind să fie mai interesaţi în grupurile mici care pot fi organizate după preferinţele sociometrice exprimate [18].

Scanned with OKEN Scanner

Studierea grupurilor umane restrînse a stat în atenţia lui J. L. M o r e n o, autorul unor metode de cercetare a configuraţiilor sociale, de care se ocupă sociometria [19], [24].

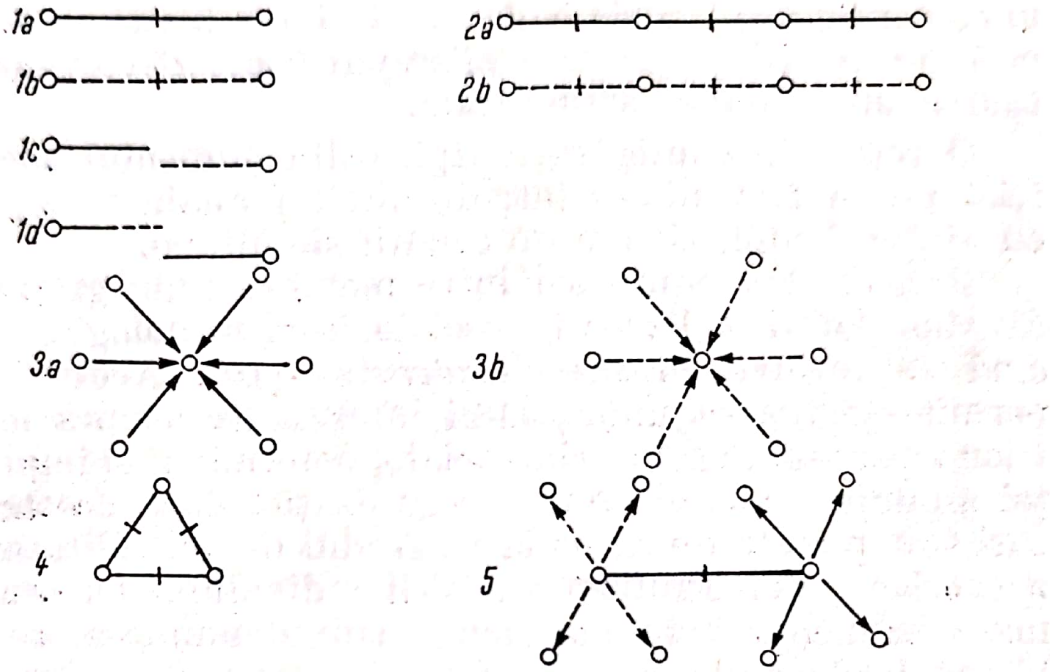


Fig. 24. — Configuraţii tipice

1 *a* — prima pereche; alegeri reciproce; 1 *b* — prima pereche; respingeri; 1 *c* — prima pereche; alegere, respingere; 2 *a* — lanţ; alegeri reciproce; 2 *b* — lanţ; respingeri reciproce; 3 *a* — stea; centrul alegerii; 3 *b* — lanţ; respingeri reciproce; 4 — triunghi; alegere reciprocă; 5 — pereche izolată; alegere reciprocă.

Aceste metode însumează mai multe procedee matematice şi grafice care prelucrează aceleaşi date oferite de testul sociometric [18].

Configuraţiile tipice sînt forme de mare universalitate, pentru că la oameni ele se remarcă indiferent dacă este vorba de clasă de elevi, grădiniţă de copii, atelier de prelucrare mecanică, birou administrativ sau fermă agricolă.

Elementele constitutive ale sociogramelor grupurilor umane restrînse sînt perechi, lanţuri, stele etc., care formează reţele mai complexe, ca de exemplu: reţeaua relaţiilor existente între elevii unei clase de învăţămînt mediu, între salariaţii unui birou dintr-o întreprindere industrială, reţeaua relaţiilor dintr-o fermă agricolă etc. Aceste reţele reprezintă configuraţiii tipice care pot fi

reprezentate prin *grafe* (orientate sau nonorientate), care sînt modele de relaţii diferite. Reprezentarea grafică a reţelelor de configuraţii tipice sau a sociogramelor, nu este numai un procedeu pentru a ilustra conţinutul unei configuraţii sociale date, ci şi un instrument care permite să se depăşească analiza pur cantitativă prin efectuarea unei analize structurale.

O reţea de configuraţii tipice din domeniul sociometriei poate fi tradusă într-un limbaj analitic, algebric, cu ajutorul unui sistem de ecuaţii simultane.

Reţelele de comunicaţii între membrii unui grup uman restrîns pot fi utilizate în sociologie şi psihologie, atunci cînd se rezolvă *problema leaderului* [16]. Aceste reţele permit efectuarea unor studii interesante asupra legăturilor existente în grupurile sociale, determinarea importanţei anumitor posturi într-o organizaţie, dacă aceste posturi sînt puncte de articulaţie, gradul de centralitate şi de periferism etc. Problema găsirii individului cu cea mai mare influenţă este cunoscută sub denumirea de problema leaderului. Aceasta este persoana care ocupă un punct central: centrul unei reţele de comunicaţie [21].

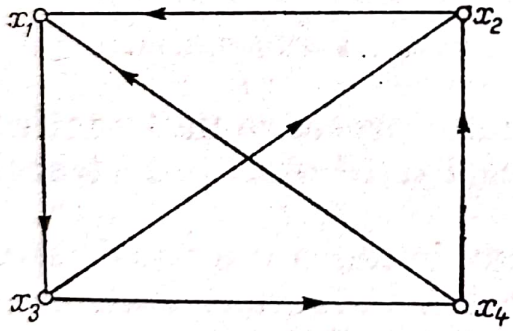


Fig. 25. — Graful unei reţele de comunicaţie între patru persoane

Vom nota cu X_i $(i = 1, 2, 3, 4)$ participanţii. Trasăm arcul $X_i\ X_j$ dacă participantul X_i îl domină pe X_j. Se obţine graful din figura 25 [16], căruia îi corespunde matricea 9.

Matricea asociată unui graf se obţine înscriind în căsuţa de la intersecţia a două rînduri numărul de arce care leagă acele vîrfuri, în sensul de parcurgere al săgeţilor, pe drumul cel mai scurt [$\delta\ (X_i, X_j) \rightarrow$ minim].

Indicele de centralitate β_i al unui nod X_i este egal cu raportul dintre suma elementelor matricei distanțelor și suma distanțelor între nodul i și celelalte noduri:

$$\beta_i = \frac{\sum_i \sum_j \delta(X_i, X_j)}{\sum_j \delta(X_i, X_j)}$$

Efectul

Cauza	x_1	x_2	x_3	x_4
x_1	0	0	1	0
x_2	1	0	0	0
x_3	0	1	0	1
x_4	1	1	0	0

Matricea 9. — Matricea asociată grafului din fig. 25

Indicele de periferism γ_i al unui nod X_i este egal cu diferența dintre indicele de centralitate al centrului grafului și indicele de centralitate al nodului X_i, adică $\gamma_i = \beta_k - \beta_i$. Modul de calcul al acestor doi indici pentru matricea 9 se dă în matricea 10 pentru calculul indicilor de centralitate și de periferism.

Se constată că nodurile X_3 și X_4 sînt amîndouă centre (matricea 10), deoarece au indicele de centralitate maxim și anume 4, 7. De asemenea și nodurile periferice au indici de periferism diferiți, punctul cel mai periferic fiind X_2, cu indicele de periferism cel mai mare:

$$\gamma_2 = \beta_{centru} - \beta_3 = 4{,}7 - 3{,}2 = 1{,}50.$$

Nodurile X_3 și X_4, avînd indicii de centralitate maximă, au indici de periferism nuli.

În graful din figura 24, persoanele X_3 şi X_4 apar ca fiind cele mai influente, deoarece ele domină persoane mai influente.

	x_1	x_2	x_3	x_4	$\sum_j \delta(X_i X_j)$	β_i	γ_i
x_1	0	2	1	2	5	$\frac{19}{5}=3{,}8$	0,9
x_2	1	0	2	3	6	3,2	1,5
x_3	2	1	0	1	4	4,7	0
x_4	1	1	2	0	4	4,7	0

$$\sum_i \sum_j \delta(X_i X_j) = 19$$

Matricea 10. — Matricea şi calculul indicilor de centralitate şi de periferism

Avantajul prezentării matriciale asupra prezentării grafice, este important : analizei grafice, prelucrarea matricială a datelor sociometrice îi substituie un proces de iteraţie, instrumentul unei analize algebrice, pe drept cuvînt sistematice [14].

Departe de a fi constrîns a-şi limita ambiţiile la o analiză de interdependenţă între membrii grupului, cercetătorul care întreprinde o investigaţie sociometrică are mijloacele necesare de a examina, în toată complexitatea lor, structurile caracteristice rezultate din împărţirea grupelor în subgrupe şi configuraţiile tipice particulare (reţeaua relaţiilor interpersonale) [15].

2.12. CONCLUZII

În ultimii ani asistăm la o extindere din ce în ce mai mare a aplicaţiilor teoriei grafelor în domeniul sociologiei, ştiinţă care acordă o importanţă crescîndă structurilor sociale. De altfel, sociologii se numără printre primii care au utilizat concepte din teoria grafelor în studiul

Scanned with OKEN Scanner

structurilor sociale, datorită conținutului extrem de bogat al acestei ramuri a teoriei mulțimilor.

Într-adevăr printr-un graf pot fi reprezentate structuri sociale foarte diferite, ca de exemplu: comunicațiile într-un grup social, circulația informațiilor într-o structură socială, evoluția populațiilor într-un fenomen demografic, relațiile de rudenie ale unui grup de indivizi, fluxurile de mobilitate geografică ale populației etc.

Aplicarea unor concepte din teoria grafelor poate contribui la aprofundarea analizei calitative a proceselor sociale, deoarece grafele utilizate în acest scop implică elaborarea unor construcții logice, simbolice și intuitive ale structurii acestor procese, concepute ca sisteme care analizează și explică întregul ansamblu de relații cauzale dintre variabile.

Probabil că, în domeniul sociologiei, teoria grafelor va aduce în viitor un aport important la formularea unor modele cauzale, adică la construirea unor asemenea reprezentări care ar permite analiza și interpretarea pe baze științifice a structurilor unor procese sociale deosebit de complexe. Utilizarea unor astfel de reprezentări ar putea deschide un cîmp larg de aplicare a matematicii calitative în acest domeniu [16].

Fără îndoială că, reprezentarea unor structuri sociale prin grafuri, alcătuite din noduri (puncte) și arce, permite efectuarea unei analize mai profunde a unui grup uman sau a unui proces social oarecare, deoarece, prin descoperirea înlănțuirii cauzale a elementelor unei astfel de structuri, putem descifra mai adînc acest grup sau acest proces și putem descoperi legile care-l guvernează. Graful legăturilor cauzale este o construcție intuitivă, simbolică, adică o rețea alcătuită din indicatori reprezentați prin noduri (puncte) și relații cauzale puse în evidență prin arce orientate. De aceea, se poate admite că graful legăturilor cauzale existente în structura unui proces social oarecare reprezintă un *model cauzal* [17].

Interesul crescînd pentru utilizarea grafelor în studiul structurilor sociale poate fi interpretat prin prisma

Scanned with OKEN Scanner

principalului avantaj oferit de această teorie, care constă în posibilitatea utilizării unor variabile nonnumerice, ordinale sau cardinale în analiza sociologică. Utilizarea acestor variabile în studiul structurilor sociale poate lărgi sfera de aplicare a matematicilor calitative în etapa de emitere a ipotezelor şi a matematicilor cantitative, în faza de verificare a acestora ; prin urmare teoria grafelor poate fi utilizată în toate fazele cercetării sociologice. Din acest punct de vedere, modelele cauzale pot fi prezentate în două moduri :

— sub formă de modele simbolice — intuitive ;
— sub formă de modele matematice.

Întrucît modelele simbolice-intuitive ale relaţiilor cauzale sînt fundamentate de regulă pe variabile nonnumerice, ele servesc mai mult înţelegerii şi numai în măsură limitată cunoaşterii structurilor analizate. Cu toate acestea astfel de modele pot marca trecerea de la modelele de înţelegere, la modelele de cunoaştere ştiinţifică. Pentru ca modelele simbolice-intuitive, care ajută la înţelegerea proceselor sociale analizate, să devină modele matematice ale relaţiilor cauzale, care să contribuie efectiv la lărgirea orizontului de cunoaştere ştiinţifică, indicatorii utilizaţi vor fi cuantificaţi, iar corelaţiile statistice şi relaţiile cauzale vor fi măsurate cu ajutorul unor metode statistico-matematice mai perfecţionate.

Cînd dorim să efectuăm o analiză matematică a unei structuri sociale, urmărim să elaborăm un model aproximativ al realităţii studiate ; acest model poate fi numeric sau nu.

Dacă în studiul unei structuri sociale, raţionăm logic, este util şi totdeauna posibil de a da o formă matematică raţionamentului nostru. Dacă aplicăm riguros această idee, matematica aplicată în domeniul sociologiei conduce la un model numeric impus în cercetarea noastră. În felul acesta, structura socială concretă este repezentată printr-o structură matematică foarte bogată, care posedă un grup numeros de proprietăţi particulare. Aceasta înseamnă să dai realităţii un model numeric, acceptînd ideea că structura socială concretă este aproximativ

Scanned with OKEN Scanner

izomorfă cu o structură numerică şi posedă proprietăţi ale acestei structuri, ceea ce este un adevăr rar. De aceea, de multe ori trebuie să recurgem la structuri nu numai numerice, ci şi la structuri nonnumerice particulare, cu condiţia ca acestea din urmă să fie structuri matematice apte să evidenţieze cîteva proprietăţi ale realităţii studiate [18], [19].

Teoria grafelor ne oferă unul din mijloacele remarcabile de analiză a structurilor sociale, deoarece este în fapt teoria relaţiilor oarecare. Dar conceptele acestei teorii, elaborate pînă în prezent, sînt insuficiente pentru a explica integral fenomenele sociale. De aceea, teoria grafelor va trebui să fie îmbogăţită cu noi concepte care să sporească eficacitatea aplicaţiilor acestei teorii în studiul structurilor sociale.

BIBLIOGRAFIE

1. Miron Constantinescu, *Aplicarea teoriei grafelor la studierea procesului de urbanizare*, în: id., *Cercetări sociologice 1938—1971*.
2. Henri H. Stahl, *Unealta de lucru a „modelului" sociologic creat de Dimitrie Gusti*, Sociologia militans, vol. II — Metode şi tehnici sociologice, 1969.
3. Henri H. Stahl, *Tehnica monografiei sociologice*, 1934. Bucureşti, Institutul social român.
4. M. Vasilescu, M. Şileţchi, L. Curta şi A. Almăşan-Radu, *Aplicaţiile algebrei matriciale în studiul structurii proceselor sociale*, în: Revista de studii şi cercetări de calcul economic şi cibernetică economică, nr. 1, 1969.
5. Ştefan Novac, *Indicatori inferenţiali, definitorii şi corelaţionali în teoria şi cercetarea sociologică*, 1963.
6. Claude Flament, *Théorie des graphes et structures sociales*, Paris, Mouton, Gauthier-Villars, 1965.
7. Claude Berge, *Théorie des graphes et ses applications*, Paris, Dunod, 1958.
8. A. G. Aganbeghian, *Kolicestvenie metodî v soţiologii*, Izd. „Nauka", 1966.
9. Anna Pawelczjnska, *The Dynamics of Cultural Changes in the Rural Areas*, în: The Polish Sociological Bulletin, nr. 1, 1966.

Scanned with OKEN Scanner

10. Oystein Ore, *Grafele și aplicațiile lor* (traducere), Edit. științifică, București, 1968.
11. Barat Maré, *Sur le mot et le concept de modèle*, în: Annales, XVIII, 1963.
12. Rudolf Carnap, *Introduction to symbolic logic and its applications*, New York, Dover, 1968.
13. Cherry Colin, *On human comunication*, New York, Wiley — Londres, Chapman and Hall, 1957.
14. Hubert M. Jr Blallock, *Further observations on asymmetric causal models*, în: American sociological review, 27, 1962.
15. Hubert M. Jr. Blallock, *Causal inferences in non-experimental research*, Chapel Hill, University of North Carolina Press, 1964.
16. Jean Cuisenier, *Economie et sociomètrie*, în: L'économique et les sciences humaines, tome I, Guy Palmade, Paris, Dunod, 1967.
17. J. L. Moreno, *Les fondements de la sociométrie* (traduit de l'anglais, d'après la 2e édition de *Who shall survive?*, Beacon House, New York, 1953), Paris, P. U. F., 1954.
18. Achim Mihu, *Sociometria, eseu critic*, „Biblioteca de filozofie și sociologie", București, Edit. politică, 1967.
19. „Statistics of social configuration", în: *The Sociometry Reader*, The Free Press of Glencoe, Illinios (U. S. A.), 1960.
20. Manea Mănescu, Anatol Măcriș, Vincențiu Dumitru, *Metode matematice moderne aplicate în organizarea și planificarea lucrărilor de construcții și montaje*, în: Manea Mănescu ș. a., *Calcul economic*, București, Edit. Academiei, 1969.
21. Luce R. Duncan și A. D. Perra, *A method of matrix analysis of group structure*, în: Psychometrica, 14, 1949.
22. A. Niceforo, *Le profil graphique des individes et des groupes, Actualités scientifiques et industrielles*, Paris, Herman, 1936.
23. Alain Badiou, *Le concept de modèle*, Paris, edit. François Maspero, 1969 (Cours de philosophie pour scientifiques 1967—1968, fascicule IV).
24. Jean Viet, *Les méthodes structuralistes dans les sciences sociales*, Paris, Edit. Mouton, 1969.
25. Achim Mihu, *Sociologia americană a grupurilor mici*, București, Edit. politică, 1971.
26. Tadeusz Pawlowski, *On the Relation between Structure of Quantitative Statements and the Types of Scales Used to Measure the Quantitation, în:* The Polish Sociological Buletin, nr. 1, 1964.

Scanned with OKEN Scanner

3. MODELE DE ANALIZĂ CORELAȚIONALĂ

3.1. INTRODUCERE

Analiza corelațională folosită în cercetările sociologice constituie o metodă de cercetare a realității sociale, bazată pe descompunerea unui întreg (fenomen, proces social complex sau structură reală) în elementele lui componente (procese sociale particulare, factori, indicatori etc.) și pe studierea relațiilor statistice dintre ele. Fiind o metodă cantitativă de analiză sociologică, cu ajutorul ei putem aprecia în ce măsură există o relație statistică între două sau mai multe variabile, care de fapt reprezintă indicatori ce descriu procese sociale particulare sau factori sociali. În asemenea cazuri vorbim de corelații dintre indicatori a căror dependență reciprocă este numai parțială. Prin urmare, putem spune că indicatorii sînt legați stocastic sau sînt corelați, în loc de a fi legați funcțional. Relațiile stocastice pot fi liniare sau neliniare, în funcție de dispersia datelor obținute în urma efectuării investigațiilor de teren. Acest ansamblu de date ne oferă posibilitatea să stabilim corelațiile cele mai probabile între indicatorii analizați.

Corelația dintre probabilitate și realitate în studiul fenomenelor sociale este dată de stabilitatea acestora într-un număr mare de investigații de teren. Această stabilitate a fenomenelor sociale constituie premisa de bază a apariției legilor statistice. Pentru a fi descoperite aceste legi, statistica matematică a elaborat mai multe metode de cercetare, care permit, pe baza datelor din investiga-

Scanned with OKEN Scanner

3. MODELE DE ANALIZĂ CORELAȚIONALĂ

3.1. INTRODUCERE

Analiza corelațională folosită în cercetările sociologice constituie o metodă de cercetare a realității sociale, bazată pe descompunerea unui întreg (fenomen, proces social complex sau structură reală) în elementele lui componente (procese sociale particulare, factori, indicatori etc.) și pe studierea relațiilor statistice dintre ele. Fiind o metodă cantitativă de analiză sociologică, cu ajutorul ei putem aprecia în ce măsură există o relație statistică între două sau mai multe variabile, care de fapt reprezintă indicatori ce descriu procese sociale particulare sau factori sociali. În asemenea cazuri vorbim de corelații dintre indicatori a căror dependență reciprocă este numai parțială. Prin urmare, putem spune că indicatorii sînt legați stocastic sau sînt corelați, în loc de a fi legați funcțional. Relațiile stocastice pot fi liniare sau neliniare, în funcție de dispersia datelor obținute în urma efectuării investigațiilor de teren. Acest ansamblu de date ne oferă posibilitatea să stabilim corelațiile cele mai probabile între indicatorii analizați.

Corelația dintre probabilitate și realitate în studiul fenomenelor sociale este dată de stabilitatea acestora într-un număr mare de investigații de teren. Această stabilitate a fenomenelor sociale constituie premisa de bază a apariției legilor statistice. Pentru a fi descoperite aceste legi, statistica matematică a elaborat mai multe metode de cercetare, care permit, pe baza datelor din investiga-

Scanned with OKEN Scanner

ţiile de teren, formularea unor concluzii fundamentate din punct de vedere ştiinţific, asupra fenomenelor sau proceselor sociale care se reproduc identic în cadrul aceluiaşi complex de condiţii. În contextul acestor metode se înscrie analiza corelaţională, care în domeniul cercetărilor sociologice poate contribui la lărgirea orizontului cunoaşterii fenomenelor sau proceselor sociale complexe, printr-o mai completă interpretare a datelor rezultate din investigaţiile de teren. În acelaşi timp, dat fiind faptul că metoda de analiză corelaţională are o bază teoretică în continuă perfecţionare, aceasta îi oferă perspective promiţătoare de dezvoltare. Astfel, în contextul acestei baze un loc important îl ocupă *teoria corelaţiei statistice.*

Analiza corelaţională folosită în cercetările sociologice constituie o metodă importantă de analiză cantitativă a relaţiilor statistice existente între factorii sau procesele sociale particulare din care este alcătuit un fenomen sau proces social complex. Utilizarea acestei metode în domeniul sociologiei asigură o reflectare mai precisă a laturilor calitative ale fenomenelor sau proceselor sociale, dat fiind faptul că deosebirile calitative au într-un fel sau altul, un suport cantitativ. Dar, pentru a fi posibil acest lucru, în prealabil este necesar să se efectueze o analiză profundă a relaţiilor statistice existente între variabilele care descriu o anumită structură socială. Efectuarea analizei calitative înaintea celei cantitative este un principiu de bază în aplicarea acestei metode la problemele concrete de sociologie. Avînd în vedere procedeele moderne de analiză corelaţională, respectarea acestui principiu nu mai poate constitui un impediment. De exemplu, utilizarea conceptelor din teoria grafelor permite efectuarea unui studiu calitativ prealabil în faza de emitere a ipotezelor.

Sarcina analizei corelaţionale constă în elaborarea procedeelor care să permită descoperirea trăsăturilor fundamentale ale fenomenului sau procesului social cercetat. În acest scop se foloseşte un volum de informaţii de obicei restrîns, rezultat în urma efectuării investigaţiilor de teren, şi care prezintă un interes practic imediat, de-

Scanned with OKEN Scanner

oarece el furnizează informațiile necesare asupra fenomenului sau procesului social studiat, care pot fi valorificate într-o măsură însemnată cu ocazia efectuării analizei corelaționale. Cu toate acestea, concluziile cele mai juste le putem trage atunci cînd aceste date se referă la un material cît mai omogen și cît mai bogat. Faptul acesta face dificilă oarecum aplicarea metodei de analiză corelațională în studiul fenomenelor sau proceselor sociale mai complexe, cum sînt cele ale vieții sociale. Aceste greutăți pot fi însă învinse prin asigurarea reprezentativității datelor culese care conferă posibilitatea obținerii unor concluzii perfect valabile. În asemenea condiții, datele cifrice pot reflecta just însușirile fenomenelor sociale studiate și pot asigura degajarea aspectelor calitative esențiale, bineînțeles cu condiția de a fi stabilite și adoptate criteriile cele mai potrivite în procesul de analiză cantitativă. În plus, metodele de investigare sociologică vor trebui astfel concepute, încît să permită valorificarea deplină a informațiilor culese cu ocazia efectuării cercetărilor de teren.

În studiul proceselor sociale observăm o multitudine de corelații statistice dintre variabile. Pentru a se asigura deducții suficient de întemeiate pe baza calculelor de corelație multiplă este necesară includerea în cercetare, în măsura în care este posibil, a tuturor factorilor esențiali. Nerealizarea acestei condiții diminuează valoarea cognitivă a indicatorilor de corelație multiplă.

Cea mai potrivită metodă pentru aprecierea importanței fiecărui factor din mulțimea factorilor corelați este analiza dispersională. Această tehnică utilizată de statistica matematică permite verificarea obiectivității legăturii observate între mai multe variabile.

Întrucît procesele sociale se produc sub acțiunea unor cauze multiple și extrem de complexe, la cercetarea legăturii acestora se includ doar factorii esențiali. Ca atare, indicatorii de corelație multiplă dau o expresie incompletă relației analizate. Într-o structură socială, mărimea unei variabile dependente este generată de acțiunea comună a variabilelor independente, considerîndu-se că acestea

din urmă alcătuiesc un sistem complet. Dacă însă luăm în calculul corelației numai o parte, rezultatele vor fi afectate de erori. De altfel, aceasta este situația obișnuită, întîlnită frecvent în practică.

O serie de erori în calculul corelației multiple se datoresc utilizării unor metode selective de culegere a informațiilor sociologice. De aceea, este absolut necesară *aprecierea probabilistică* a indicatorilor corelației multiple.

Aplicarea metodei corelației multiple în analiza gradului de tărie al tuturor relațiilor statistice dintre caracteristicile unei structuri sociale, poate fi foarte utilă într-o cercetare sociologică. De pildă, dacă avem în vedere faptul că în domeniul sociologiei cercetăm adesea o serie de structuri sociale, apare deosebit de oportună efectuarea unei analize statistico-matematice a conexiunii legăturilor dintre variabilele ce descriu aceste structuri complexe.

3.2. NATURA RELAȚIILOR STATISTICE

Aplicarea metodelor statistico-matematice în analiza proceselor sociale implică evidențierea relațiilor statistice existente între diferite variabile care descriu o serie de factori sociali sau procese sociale particulare. Pentru evidențierea acestor relații de multe ori se apelează la studiul corelației statistice, evident în scopul măsurării și exprimării intensității legăturii dintre variabile.

Structura unui proces social oarecare este deosebit de complexă, cuprinzînd un ansamblu de relații statistice, datorită faptului că această structură apare ca o împletire multiplă de cauze și efecte; de obicei observăm efecte produse de mai multe cauze, fiecare din ele contribuind într-o măsură diferită la producerea efectului. În cadrul proceselor sociale se contopește acțiunea unei multitudini de factori (economici, psihici, culturali, biologici, geografici etc.) și de relații cauzale. Factorii pot fi

Scanned with OKEN Scanner

determinanți sau determinați, iar relațiile cauzale pot fi esențiale sau neesențiale.

Datorită complexității relațiilor cauzale existente în structura socială, studiul statistic al acestora are o importanță deosebită.

Studiind relațiile statistice existente între diferitele variabile care descriu procese particulare (economice, culturale, psihice, geografice, biologice etc.), statistica matematică nu se poate limita la măsurarea acestei legături din punct de vedere cantitativ, ci trebuie să explice esența lor, calitatea lor internă. Pentru stabilirea relațiilor dintre variabilele unei structuri sociale, statistica-matematică trebuie să aibă în vedere unitatea dialectică dintre latura calitativă și latura cantitativă a proceselor sociale, adică legătura indisolubilă dintre calitatea proceselor cercetate și determinarea lor cantitativă, prin aplicarea metodelor statistico-matematice. Cu ajutorul acestor metode se poate măsura *intensitatea* acestei legături, se poate exprima forma și direcția ei.

În cadrul proceselor sociale pot exista o varietate de relații statistice. Stabilirea tipurilor de relații statistice se face în urma unei analize cantitative multilaterale.

Sînt cazuri în care o variabilă este cauza, iar o altă variabilă este efectul; în această situație avem de-a face cu o *relație cauzală simplă*, cum ar fi de pildă creșterea nivelului de școlarizare al unei persoane care determină creșterea nivelului de calificare al ei. Alteori, între variabile există o *relație reciprocă, o interdependență* cînd nici una din variabile nu poate fi considerată cauză sau efect. Există posibilitatea ca variația a două variabile să se datoreze *influenței unei a treia variabile*, care le determină pe amîndouă.

Se întîlnesc cazuri — îndeosebi în natură și tehnică — în care o variabilă determină în mod univoc variația celeilalte variabile. Aceasta înseamnă că unei anumite valori a uneia dintre caracteristici îi corespunde numai o

Scanned with OKEN Scanner

valoare determinată a celeilalte variabile. Această situaţie o întîlnim în cazul unor relaţii deterministe, adică *relaţii de tip funcţional*. Datorită complexităţii lor, proceselor sociale cu relaţii deosebit de variate între caracteristici, nu le sînt specifice aceste relaţii funcţionale. Dependenţele unor variabile faţă de altele se manifestă numai în general, ca o tendinţă medie, într-un număr mare de cazuri. De aceea, relaţiile de acest fel se numesc *relaţii statistice* [1]. Specificul acestor relaţii constă în faptul că o variabilă influenţează o altă variabilă cu o intensitate mai mare sau mai mică; în acest caz unei valori a variabilei determinante îi corespund mai multe valori ale variabilei determinate. În acelaşi timp, o variabilă poate fi influenţată concomitent de mai multe variabile. De pildă, dacă notăm variabila determinată cu y, iar caracteristicile care o influenţează cu x_j, vom avea:

$$y = f(x_1, x_2, \ldots, x_n).$$

De exemplu, între productivitatea muncii şi o serie de factori economici şi sociali există o relaţie statistică de necontestat, deoarece prima variabilă depinde de înzestrarea tehnică, de calificarea muncitorilor, organizarea locului de muncă, de sex, de vîrstă etc.

Prin urmare, pentru procesele sociale sînt specifice relaţiile statistice, care sînt studiate cu ajutorul metodelor statistico-matematice. Cu ajutorul metodelor sale, statistica descoperă ceea ce este necesar din multitudinea întîmplărilor şi dă o expresie numerică dezvoltării proceselor sociale.

Specificul corelaţiilor statistice constă în faptul că acestea nu pun în evidenţă relaţii cauzale între variabile, ci o variaţie concomitentă a lor. De aceea, corelaţiile statistice pot semnala cel mult o asociere demonstrabilă între variaţia a două sau mai multe variabile.

Scanned with OKEN Scanner

3.3. CLASIFICAREA RELAȚIILOR STATISTICE

Relațiile statistice pot fi clasificate după mai multe criterii: după numărul de caracteristici (variabile), după modul de exprimare al caracteristicii, după direcția relației și după expresia analitică a ei.

După *numărul de caracteristici*, a căror relație se cercetează, putem avea *relații simple*, cînd luăm în studiu numai două variabile și *relații multiple*, cînd studiem mai mult de două variabile. Un exemplu de relație simplă ar fi relația între nivelul de școlarizare al subiectului și gradul de calificare al acestuia, iar un exemplu de relație multiplă ar fi relația între coeficientul de migrație al populației de la sat la oraș, gradul de urbanizare și gradul de industrializare al zonei orășenești.

După *modul de exprimare al caracteristicii* există relații dintre *caracteristici exprimate numeric*, ca de exemplu relația între productivitatea muncii și categoria de încadrare; relații între *o caracteristică exprimată numeric, cantitativ și alta exprimată calitativ*, ca de pildă relația între profesiune și salariu; în fine relații între două *caracteristici exprimate calitativ*, ca de pildă relația între caracteristici exprimate prin cuvinte: profesiune și studii.

După *direcția relației statistice* care se stabilește între variabilele analizate există *relație directă*, atunci cînd creșterea caracteristicii determinante (x) duce la creșterea caracteristicii determinate (y) și *relație inversă*, cînd creșterea caracteristicii determinante (x) duce la descreșterea caracteristicii determinate (y). Un exemplu de relație directă este cea între înzestrarea tehnică și creșterea productivității muncii, iar un exemplu de relație indirectă este cea între gradul de fluctuație al muncitorilor și productivitatea muncii.

După *expresia analitică* a legăturii statistice care se stabilește, relația poate fi liniară, cînd se exprimă prin

Scanned with OKEN Scanner

ecuaţia liniei drepte, şi curbilinie cînd se exprimă prin ecuaţia unei curbe (parabolă, hiperbolă, funcţie exponenţială).

3.4. PRINCIPII TEORETICE ALE ANALIZEI RELAŢIILOR STATISTICE

În cazul proceselor sociale stabilirea şi măsurarea anumitor relaţii statistice între diferite variabile devin mai dificile datorită complexităţii şi multitudinii factorilor care acţionează într-o îmbinare permanentă, fiecare influenţînd cu o intensitate mai mare sau mai mică, direct sau indirect. Din această cauză, metodele statistice de stabilire şi de măsurare a acestor relaţii trebuie să fie aplicate cu foarte multă atenţie, pentru a nu fi golite de conţinut. De aceea este necesar ca, înainte de aplicarea unei metode statistice de stabilire a relaţiilor, să avem în vedere respectarea cîtorva principii teoretice în cercetarea proceselor sociale.

Aplicarea cu succes a metodei corelaţiei este condiţionată şi trebuie să fie precedată de o analiză calitativă multilaterală a proceselor sociale cercetate, desfăşurată în următoarele direcţii [2]:

a) *identificarea existenţei unei relaţii* între variabilele analizate. Numai în cazul în care există relaţii cauzale reale, evidente, stabilite de sociologie ca ştiinţă, se va trece la verificarea lor statistică;

b) necesitatea ca în analiza anumitor relaţii statistice să ţinem seama şi *să includem în calcul toţi factorii determinanţi* care acţionează sistematic asupra factorului-efect, adică să calculăm o corelaţie multiplă şi apoi să efectuăm o analiză cauzală.

În practică se întîlnesc însă situaţii în care dorim să măsurăm relaţia statistică între două variabile. În acest caz avem de-a face cu o corelaţie simplă, în care o variabilă determinantă sau factorială (x) influenţează o altă

Scanned with OKEN Scanner

variabilă determinată sau rezultativă (y). Determinarea unei relații statistice dintre două variabile este posibilă, dar în acest caz trebuie să facem abstracție de ceilalți factori chiar dacă sînt esențiali, considerîndu-i neesențiali și întîmplători față de factorul-cauză cercetat;

c) desprinderea *formei relației statistice* dintre variabile. Pentru stabilirea existenței și formei acestei relații, a tendinței și a intensității ei, statistica dispune de o serie de metode, unele mai simple altele mai complicate;

d) *stabilirea unor relații statistice, a unor corelații între variabile prin interpretarea informațiilor obținute dintr-o observare totală sau dintr-o observare selectivă* [3]. În cazul unei observări selective, indicatorul corelației statistice se apreciază cu o anumită eroare, care trebuie să fie cuprinsă în limite acceptabile.

3.5. METODE ELEMENTARE DE CALCUL A CORELAȚIILOR SIMPLE

Corelațiile simple pot fi stabilite cu ajutorul unor metode elementare, cum ar fi: metoda seriilor paralele interdependente, metoda grupărilor, metoda tabelelor de corelație, metoda grafică etc. [1].

Metoda seriilor paralele interdependente constă în stabilirea corelațiilor simple prin compararea a două serii de termeni. În acest caz, variabilele care se referă la caracteristica factorială se așează în ordine crescătoare sau descrescătoare, iar valorile variabilelor rezultative se înscriu paralel cu prima serie. De pildă, în acest fel pot fi comparate diferite serii cronologice, teritoriale și de distribuție. Aplicarea acestei metode dă posibilitatea stabilirii existenței și direcției de dezvoltare a relației statistice fără a se putea aprecia gradul de intensitate al acestei relații. De exemplu, în tabelul 7 se dau două serii paralele cronologice privind dinamica fondurilor fixe productive (x) și producti-

Scanned with OKEN Scanner

vitatea muncii pe salariat în industrie (y) în procente, pe perioada 1960 — 1966 (anul de bază este considerat 1950 = 100).

Tabelul 7

Anii	Fondurile fixe productive în industrie (în %) (x)	Productivitatea muncii pe salariat în industrie (y)
1960	225	237
1961	243	253
1962	263	270
1963	299	291
1964	327	320
1965	357	343
1966	404	372

Sursa : Anuarul statistic al R. S. România, 1967, p. 106 și 112.

Comparînd cele două serii paralele constatăm că între ele există o relație directă, deoarece creșterea fondurilor fixe productive (x) determină o creștere a productivității muncii pe salariat în industrie (y).

Metoda grupărilor statistice permite stabilirea existenței și direcției relației dintre două caracteristici prin gruparea unităților colectivității studiate după caracteristica factorială, iar pentru caracteristica rezultativă prin calcularea indicatorilor derivați, mai ales mărimi medii și mărimi relative pe grupele în care a fost împărțită colectivitatea respectivă. Atunci cînd gruparea se face numai după o caracteristică factorială și alta rezultativă avem o *grupare simplă*. Dacă însă avem mai multe caracteristici factoriale și o singură caracteristică rezultativă avem o *grupare combinată*. Din combinarea metodei grupărilor cu metoda mediilor calculate pentru caracteristica rezultativă obținem mediile pe grupă.

Pentru exemplificarea acestei metode vom efectua o grupare simplă între caracteristicile : valoarea fondurilor fixe în I. M. A. (în milioane lei) și prețul de cost al unui

Scanned with OKEN Scanner

ha a. n. (în lei). Întrucît este mai dificil să observăm tendința generală a relației inverse dintre aceste două caracteristici, se grupează datele statistice, ca în tabelul 8.

Tabelul 8

Grupele de I.M.A. după valoarea fondurilor fixe (mil. lei)	Nr. de I.M.A. (pe grupe)	Total pe grupe (lei)	
		Preț de cost al unui ha a.n.	Preț de cost mediu pe grupe al unui ha a.n.
1	2	3	4
12 — 16	8	722,6	90,3
16 — 20	8	588,6	73,5
20 — 25	12	814,5	67,8
25 — 30	10	583,8	58,3
Peste 30	2	113,1	56,5
Total	40	2822,6	70,7

Pentru a scoate în evidență relația dintre cele două caracteristici cercetate s-a calculat, în col. 4, prețul de cost mediu pe grupe în lei (col. 3 : col. 2 = col. 4). Comparînd în acest fel valoarea fondurilor fixe (col. 1) cu prețul de cost mediu (col. 4) se poate observa o relație strînsă, inversă, între creșterea caracteristicii factoriale și descreșterea caracteristicii rezultative.

La stabilirea unei relații cu ajutorul metodei grupărilor trebuie să vedem dacă între cele două caracteristici există o relație care poate fi scoasă în evidență prin intermediul acestei metode, apoi să stabilim cît mai corect numărul de grupe și mărimea intervalului de grupare în așa fel încît să redea cît mai fidel conținutul procesului social analizat. Utilizarea mediilor pe grupă în cadrul acestor grupări elimină influența factorilor întîmplători în cea mai mare măsură și evidențiază în special influența factorilor esențiali.

În cazul prezenței unor relații între trei sau patru variabile trebuie să aplicăm o grupare combinată, care face

Scanned with OKEN Scanner

posibilă o analiză mai profundă a proceselor sociale, a gradului de influenţă pe care îl exercită o serie de factori esenţiali.

Colectivitatea, în acest caz, se desparte în grupe mai întîi după o singură caracteristică, apoi aceste grupe se împart după două caracteristici şi acestea la rîndul lor se pot despărţi în grupe mai mici după a treia caracteristică etc. De regulă, nu se recomandă o împărţire a colectivităţii după prea multe caracteristici întrucît în acest fel s-ar fărîmiţa procesul social şi s-ar pierde din vedere relaţiile esenţiale pe care trebuie să le stabilim.

Metoda tabelului de corelaţie se bazează de fapt tot pe metoda grupării fiind o formă specială a metodei amintite, aplicate unei grupări simple. În tabelul de corelaţie, gruparea se face după ambele caracteristici, atît după variaţia caracteristicii factoriale cît şi după variaţia caracteristicii rezultative. La întocmirea acestui tabel se recomandă ca intervalele de grupare ale celor două caracteristici să fie egale, iar numărul grupelor formate pe baza caracteristicilor factoriale, să fie egale cu numărul grupelor obţinute în funcţie de variaţia caracteristicii rezultative. De regulă, grupele după caracteristica factorială se trec pe orizontală, în ordine crescătoare, în capul coloanelor, iar grupele după caracteristica rezultativă pe verticală, în ordinea descrescătoare, în capul rîndurilor. Numărul unităţilor colectivităţilor cercetate (frecvenţele) se trec în căsuţele formate din întretăierea coloanelor cu rîndurile. După modul în care se distribuie frecvenţele ne dăm seama de existenţa intensităţii şi direcţiei relaţiei dintre două caracteristici studiate.

Se consideră că un tablou cu dublă intrare (fig. 26) avînd m linii şi n coloane, formează un ansamblu de elemente care sînt grupate astfel:

a) fiecărei coloane îi corespunde o valoare determinată x_i (valori izolate ale unui interval de clasă) a unei caracteristici x;

Scanned with OKEN Scanner

b) fiecărei linii îi corespunde, în mod analog, o valoare determinată y_j (izolează un interval de clasă) a unei caracteristici y;

c) fiecare căsuță indică un număr n_{ij} care corespunde simultan unui cuplu de valori x_i și y_j a două caracteristici.

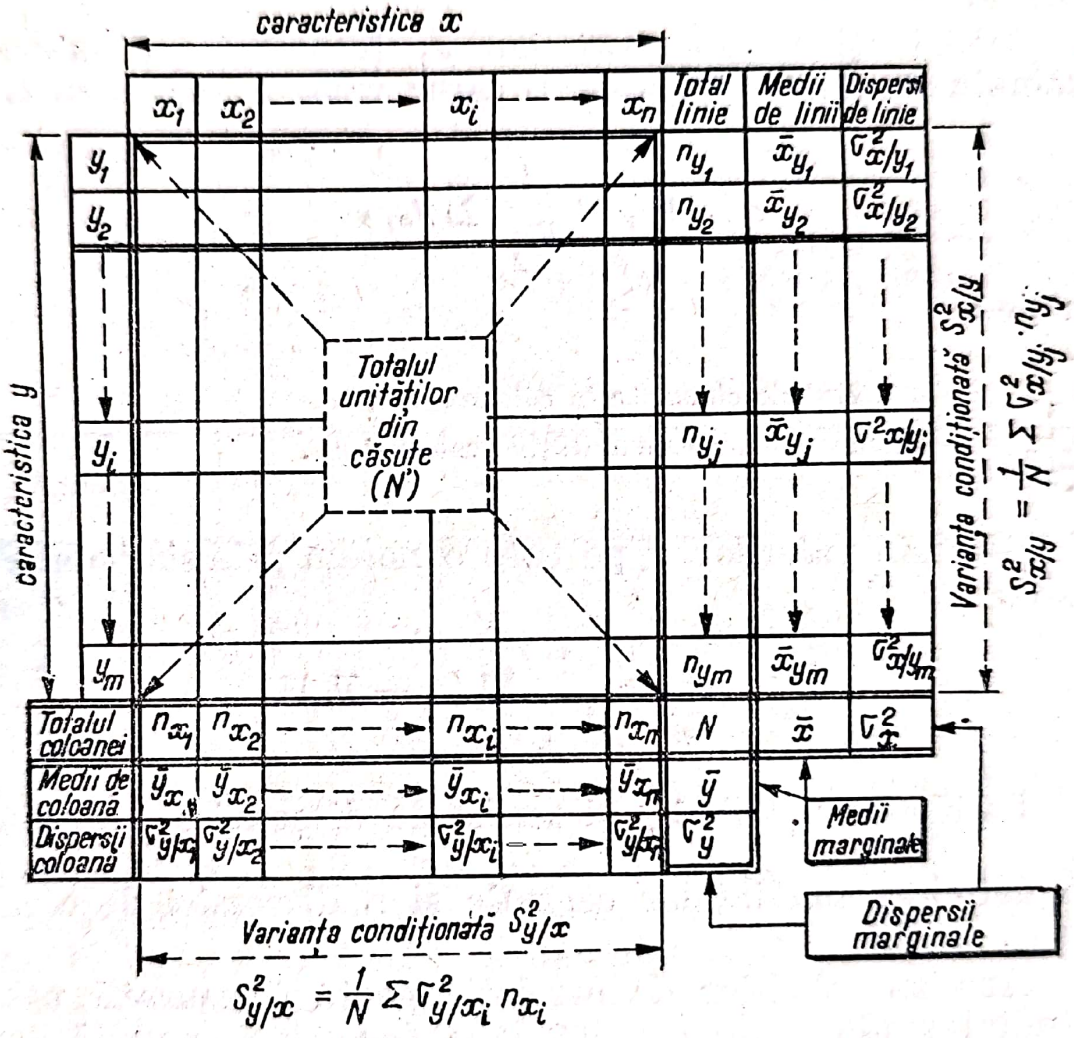


Fig. 26. — Tabelul de corelație

Cînd aceste caracteristici sînt legate între ele, această tabelă se numește *tabelă de corelație* (fig. 26) [4].

Linia de bază și coloana din dreapta figurii 26 sînt destinate totalului coloanei n_{x_i} și respectiv totalului liniei n_{y_j}, totalul unităților din căsuțele tabelei fiind N.

Valorile y_j care figurează în tabel constituie *seriile marginale*. Aceste serii sînt utilizate atît pentru determinarea mediilor ansamblului, adică a mediilor marginale $\bar{x}$ şi $\bar{y}$ ale celor două caracteristici, cît şi pentru dispersiile ansamblului, adică a dispersiilor marginale şi a dispersiilor totale, σ_{x^2} şi σ_{y^2}, care indică cazul de independenţă.

Distribuţiile caracteristicilor x_i formează un subansamblu de date statistice, a căror valoare medie va fi:

$$\bar{y}_{x_i} = \frac{1}{n_{x_i}} \sum y_{x_i},$$

unde:

n_{x_i} — numărul de elemente în coloana x_i;
$\sum y_{x_i}$ — suma valorilor y ale unităţilor coloanei x_i.

Dispersia valorilor x_i pe care o notăm prin simbolul σ^2 este:

$$\sigma^2_{\underset{x_i}{y}} = \frac{1}{n_{x_i}} \sum (y - \bar{y}_{x_i})^2.$$

Seriile de medii $\bar{x}_{y_j}$ şi seriile de dispersii $\sigma^2_{\underset{y_j}{x}}$ servesc la definirea tendinţelor centrale şi a dispersiei de corelaţie.

Analiza tabelelor de corelaţie poate fi uşurată prin studiul modului de distribuţie a frecvenţelor unităţilor colectivităţii cercetate, care ne permite să facem aprecieri asupra intensităţii şi direcţiei relaţiei dintre două caracteristici examinate. În acest scop, acelaşi tabel de corelaţie se prezintă în două moduri:

— cu totalul tuturor coloanelor considerat 100%;
— cu totalul liniilor luat 100%.

Scanned with OKEN Scanner

Spre exemplificare, se dau tabelele 9 și 10 care reflectă aceste două moduri de prezentare a datelor culese din investigații sociologice [1].

Tabelul 9

Vîrsta	Categoria de încadrare						
	1	3	4	5	6	7	8
⩽ 24 ani	29,6	30,4	19,8	18,2	10,2	—	—
25—44	62,2	58,0	62,9	65,9	74,4	62,5	83,3
45—59	8,2	11,6	17,3	15,9	15,4	37,5	16,7
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Tabelul 10

Vîrsta	Categoria de încadrare							Total
	1	3	4	5	6	7	8	
⩽ 24 ani	14,4	53,0	22,7	7,9	2,0	—	—	100
25—44	12,0	40,3	28,7	11,4	5,7	1,0	1,0	100
45—59	7,1	36,3	35,4	12,4	5,3	2,6	0,9	100

Analizînd procentele înscrise în căsuțele tabelelor 9 și 10, observăm că, cu cît trecem la grupări de vîrstă mai mari, crește proporția celor care au categorii de încadrare mai mari. De pildă, în timp ce la grupa de vîrstă pînă la 24 ani, majoritatea muncitorilor au categoria I—III, la celelalte grupe de vîrstă, ponderea celor din categoriile mai mari crește o dată cu vîrsta. În același timp, pentru categoriile I—III scade mereu ponderea celor cu vîrstă mai mare, iar la categoriile mai mari, categoria de încadrare crește cu vîrsta.

[1] Exemplul este luat din lucrarea *Analiza datelor privind unii indicatori ai forței de muncă și influența lor asupra productivității muncii*, elaborată în cadrul Centrului de cercetări sociologice (colectiv: H. Cazacu, O. Hoffman, M. Mocanu, și Gh. Socol).

Scanned with OKEN Scanner

Prin urmare, caracteristicile privind categoria de încadrare şi vîrsta muncitorului sînt corelate în mod direct : o dată cu creşterea vîrstei, creşte şi categoria de încadrare.

Metoda grafică permite prezentarea expresivă a relaţiilor existente între diferite caracteristici studiate. Această metodă are avantajul de a fi utilizată atît independent, cît şi în combinaţie cu alte metode, completîndu-se cît mai sugestiv.

Metoda grafică se bazează pe sistemul de coordonate rectangulare, în care valorile caracteristicilor factoriale (x) se trec pe axa absciselor, iar valorile caracteristicii facultative (y) se trec pe ordonată. Unităţile purtătoare ale celor două caracteristici (frecvenţele) se reprezintă printr-un punct a cărui abscisă este valoarea caracteristicii factoriale iar ordonata, valoarea caracteristicii rezultative. După modul în care sînt distribuite aceste puncte în grafic ne dăm seama de existenţa sau de absenţa unei relaţii dintre variabilele analizate.

Împrăştierea punctelor în jurul unei anumite linii, poate diferi de la o formă a relaţiei statistice la alta. Această linie poate fi dreaptă sau curbă (parabolă, hiperbolă, exponenţială etc.). Concentrarea punctelor în jurul unei drepte oarecare denotă o relaţie statistică intensă între cele două caracteristici studiate, pe cînd împrăştierea punctelor în grafic semnalează o lipsă a acestei relaţii. Reprezentarea acestor puncte în grafic furnizează o imagine a cîmpului de corelaţie, asemănătoare cu un nor de puncte. De altfel, datorită acestei asemănări, procedeul s-a denumit : *„procedeul norilor de puncte”* [5].

Cu ocazia cercetării corelaţiilor, se pot întîlni următoarele tipuri de asocieri : [4].

— corelaţii perfecte, pozitive sau negative ;
— corelaţie pozitivă sau negativă ;
— lipsa totală a corelaţiei.

Scanned with OKEN Scanner

Dacă corelațiile se pun în evidență grafic, atunci se folosesc diagrame de dispersie a datelor culese pe cupluri de caracteristici sau indicatori (x și y), care pentru situațiile menționate se vor prezenta ca în fig. 27—30.

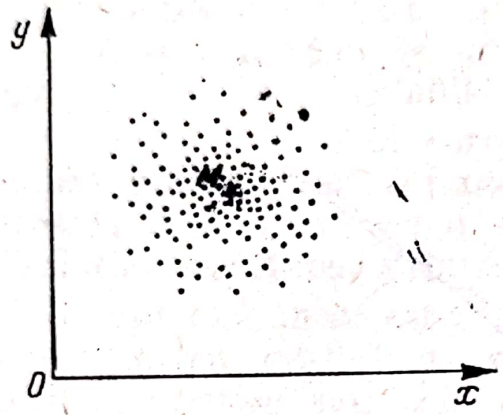


Fig. 27. — Independența între variabile

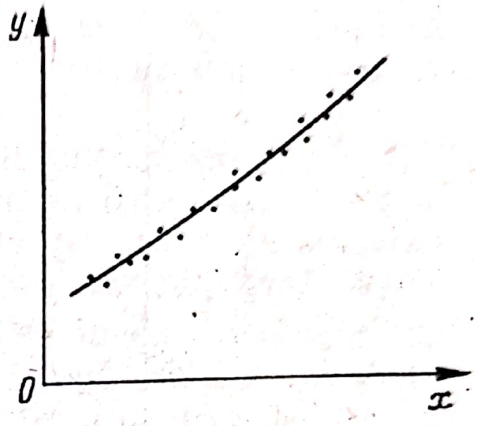


Fig. 28. — Corelație perfectă

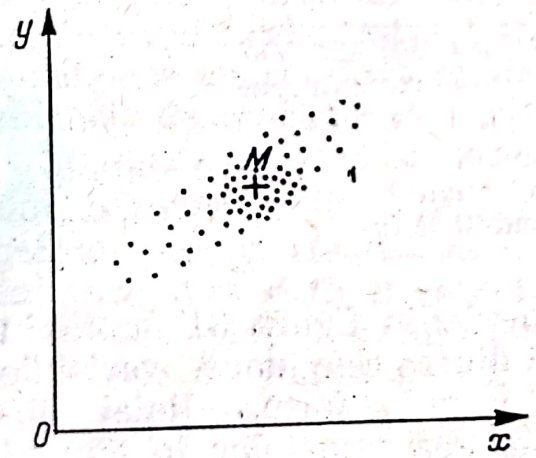


Fig. 29. — Corelație directă

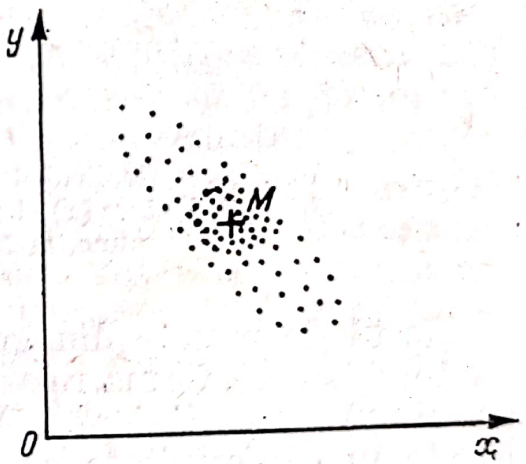


Fig. 30. — Corelație inversă

În exemplele de corelații posibile din aceste figuri, s-au reprezentat grafic norii de puncte în diagrame de dispersie, indicîndu-se și centrul de greutate corespunzător. În funcție de dispersia acestor puncte, pot exista: o independență între variabile (fig. 27), o corelație perfectă (fig. 28), o corelație directă (fig. 29) și o corelație inversă (fig. 30).

Pentru exemplificare, vom reprezenta grafic în figura 31 datele statistice care caracterizează relaţia dintre fondurile fixe productive — gradul de înzestrare tehnică a muncii — (x) şi productivitatea muncii pe muncitor în industrie (y), în perioada 1955—1965.

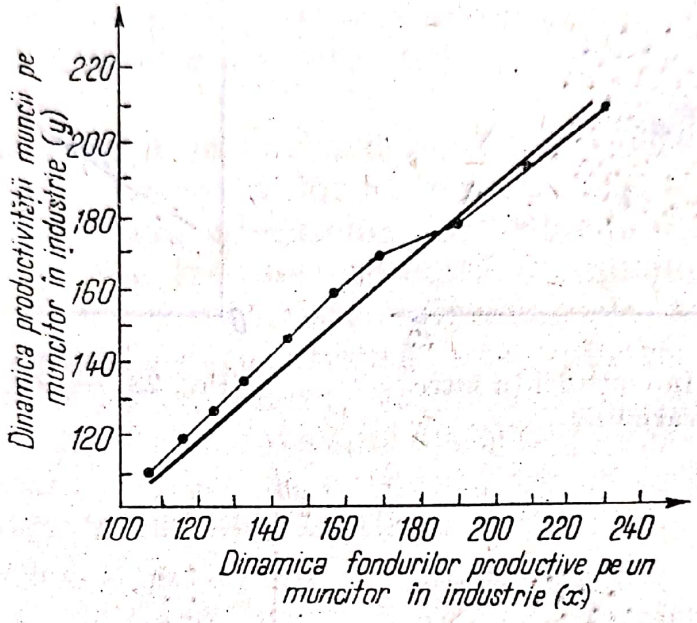


Fig. 31. — Graficul de corelaţie dintre fondurile fixe productive (x) şi productivitatea muncii pe muncitor în industrie (y)

Norul de puncte din graficul din figura 31 arată în mod evident că forma relaţiei dintre cele două variabile (x şi y) este *liniară*. De regulă, forma liniei obţinute în conformitate cu aşezarea punctelor în cîmpul graficului îl ajută pe cercetător să aleagă acea ecuaţie matematică care să exprime cît mai exact forma relaţiei.

Împrăştierea punctelor de o parte şi de alta a liniei drepte arată că asupra productivităţii muncii pe un muncitor în industrie au influenţat şi alţi factori, în afară de înzestrarea tehnică a muncii, considerată factor determinant. Influenţa factorilor întîmplători, neesenţiali se poate elimina într-o oarecare măsură prin stabilirea li-

Scanned with OKEN Scanner

niei de regresie teoretică (în cazul nostru, este o linie dreaptă).

Important în stabilirea liniei de regresie (dreaptă sau curbă) nu este alegerea unuia sau altuia dintre sistemele de funcții matematice, ci caracterul logic al relației studiate. Cu alte cuvinte, nu se poate afirma că în analiza corelației se aplică un sistem de funcții de regresie universal valabil.

Caracterul logic al corelației îl putem surprinde din analiza multilaterală a fenomenelor a căror relație se cercetează, completată cu o serie de procedee de stabilire a relației statistice, printre care și metoda grafică [6].

În practică însă, este bine să se înceapă calculul cu reprezentarea grafică a relației, după ce în prealabil datele statistice sau cele provenite din investigații au fost înscrise într-un tabel de corelație; un astfel de tabel permite stabilirea direcției și a intensității dependenței dintre două variabile. Aplicarea metodelor grafice însă ne dezvăluie relațiile statistice în toată amploarea lor, bineînțeles cu condiția ca datele culese să fie în număr suficient de mare și să manifeste regularitate.

Metodele seriilor paralele interdependente, a grupărilor, a tabelelor de corelație și a graficelor, expuse mai înainte, permit efectuarea analizei calitative a relațiilor statistice, fără să fie necesară măsurarea gradului de tărie al acestor relații. Datorită acestui fapt, aceste metode au o serie de limite ce nu pot fi trecute cu vederea în toate cazurile întîlnite în practică [7].

3.6. METODE ANALITICE DE CALCUL A CORELAȚIILOR SIMPLE

Metodele elementare de calcul a relațiilor dintre variabile nu sînt suficiente pentru o cercetare mai profundă a structurilor sociale. Aceste metode nu pot să stabi-

lească, să măsoare, prin intermediul unui indicator, intensitatea unei relaţii între două sau mai multe variabile. Pentru stabilirea unor astfel de indicatori sintetici, trebuie să apelăm la metodele analitice de cercetare a corelaţiei, care oferă posibilitatea de a aprofunda esenţa proceselor sociale, de a măsura relaţiile statistice şi de a exprima aceste relaţii printr-un indicator numeric [8].

Pentru calculul acestor indicatori ai corelaţiei trebuie să avem în vedere aspectele analizei calitative care trebuie să preceadă întotdeauna o cercetare statistică, fără de care stabilirea unei corelaţii poate să devină un exerciţiu formal.

Indicatorii cei mai utilizaţi pentru exprimarea sintetică a relaţiei dintre două sau mai multe variabile sînt : *coeficientul de corelaţie, raportul unic de corelaţie şi coeficientul de corelaţie reciprocă* [9].

Valoarea *coeficientului de corelaţie* variază între —1 şi 1. Cu cît corelaţia este mai intensă cu atît valorile coeficentului de corelaţie se vor apropia de 1, în cazul unei corelaţii directe, şi de —1 în cazul unei corelaţii inverse. Lipsa unei corelaţii indică un coeficient de corelaţie zero sau chiar apropiat de zero.

Între procesele sociale pot fi stabilite atît relaţii statistice liniare cît şi neliniare. Relaţiile neliniare pot fi reprezentate grafic cu ajutorul unor curbe : parabolă, hiperbolă, funcţie exponenţială etc.

Pentru măsurarea intensităţii relaţiei dintre două variabile se utilizează în anumite cazuri şi *raportul unic de corelaţie* care poate lua valori între 0 şi 1. În acest caz, corelaţia va fi cu atît mai intensă cu cît variaţia sistematică a caracteristicii rezultative (y), provocată de modificarea caracteristicii factoriale (x), va fi mai mare, iar raportul unic de corelaţie se apropie de 1. Dacă influenţa factorilor întîmplători este mai mare, intensitatea relaţiei statistice este mai slabă şi raportul unic de corelaţie se apropie de zero.

Scanned with OKEN Scanner

Cînd raportul unic de corelație este egal cu 1, între caracteristicile studiate există o relație funcțională; iar în cazul în care acest raport este egal cu zero, între variabilele respective nu există nici un fel de relație statistică. În cazul utilizării coeficientului de corelație reciprocă, acesta poate lua valori între 0 și 1. Sub raportul valorilor pe care le poate lua, acest coeficient devine interpretabil în același mod ca și raportul unic de corelație [10]. De exemplu, indicele de corelație liniară între venitul mediu pe o persoană și cheltuielile medii lunare pentru alimente, pe o persoană, va reflecta o corelație dintre doi indicatori cantitativi; raportul unic de corelație între bugetul zilnic de timp liber și sexul persoanei anchetate va oglindi o corelație între un indicator cantitativ și un indicator calitativ; coeficientul de corelație liniară reciprocă între dorința de a lucra la un utilaj automat și sexul persoanei anchetate va caracteriza o corelație între doi indicatori calitativi.

Acum să exemplificăm calculul acestor indicatori:

Coeficientul de corelație liniară între venitul mediu pe o persoană și cheltuielile medii lunare pentru alimente se poate determina cu ajutorul formulei [11]:

$$r = \frac{\sum (x - \overline{x})(y - \overline{y})}{\sqrt{\sum (x - \overline{x})^2 \sum (y - \overline{y})^2}},$$

în care:

x — veniturile medii lunare pe o persoană, exprimate în lei;
y — cheltuielile medii lunare pe o persoană pentru alimente, exprimate în lei.

În acest caz, bazîndu-ne pe datele din tabelul 11, procedeul de calcul va fi următorul [12]:

Scanned with OKEN Scanner

Tabelul 11

Veniturile medii lunare pe o persoană (x)	Cheltuielile medii lunare pentru alimente pe o persoană (y)	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$
300	190	−60	−50	+3000	3600	2500
320	200	−40	−40	+1600	1600	1600
340	230	−20	−10	+ 200	400	100
360	240	—	—	—	—	—
380	250	+20	+10	+ 200	400	100
400	270	+40	+30	+1200	1600	900
420	300	+60	+60	+3600	3600	3600
2520	1680	—	—	9800	11200	8800

$$\bar{x} = \frac{2520}{7} = 360; \qquad \bar{y} = \frac{1680}{7} = 240;$$

$$r = \frac{9800}{\sqrt{11200 \times 8800}} = \frac{9800}{\sqrt{98560000}} = \frac{9800}{9928} = 0{,}987.$$

Raportul unic de corelaţie între bugetul zilnic de timp liber şi starea civilă a persoanei anchetate se poate determina cu ajutorul formulei [13]:

$$\eta = \sqrt{\frac{\sum n_j (\bar{y}_x - y)^2}{\sum\sum n_{ij} (y - \bar{y})^2}},$$

în care:

$\bar{y}_x$ — media coloanei din tabela de contingenţă;
$\bar{y}$ — media generală;
n_j — efectivul coloanei din tabela de contingenţă;
n_{ij} — efectivul căsuţei respective din tabela de contingenţă;
$\Sigma\Sigma$ — indică faptul că se totalizează efectivele coloanelor.

Scanned with OKEN Scanner

În acest caz, bazîndu-ne pe datele din tabelul 12, procedeul de calcul va fi următorul:

Tabelul 12

Bugetul zilnic de timp liber	Sexul persoanei anchetate		Total
	masculin	feminin	
—	6	15	21
0,30	1	1	2
0,50	2	1	3
0,70	1	1	2
0,80	—	3	3
1,00	—	2	2
1,70	1	—	1
2,00	2	—	2
2,50	2	1	3
2,70	—	1	1
3,00	1	—	1
3,20	1	—	1
3,40	1	—	1
3,50	2	—	2
4,00	1	—	1
5,40	—	1	1
5,70	1	—	1
6,20	—	1	1
8,00	1	—	1
Total	23	27	50

$$\eta = \sqrt{\frac{23(2,04 - 1,40)^2 + 27\,(0,84 - 1,40)^2}{139,83}} =$$

$$= \sqrt{\frac{17,80}{39,83}} = 0,36 \cdot$$

Coeficientul de corelație reciprocă (Pearson) între dorința de a lucra la un utilaj automat și sexul persoanei anchetate

se poate determina cu ajutorul formulei [14]:

$$c = \sqrt{1 - \frac{1}{\sum\sum \frac{n_{j_1 j_2}^2}{n_{j_1} n_{j_2}}}},$$

în care:

n_{j1} — frecvența unei grupe de indicatori studiați;
n_{j2} — frecvența altei grupe de indicatori studiați;
n_{j2}^{j1} — frecvența combinațiilor corespunzătoare ale acestor indicatori.

În acest caz, bazîndu-ne pe datele din tabelul 13, procedeul de calcul va fi următorul:

Tabelul 13

Nr. crt.	Dorința de a lucra la un utilaj automat (n_{j1})	Sexul persoanei anchetate (n_{j2})		Total
		Masculin	Feminin	
1	Ar dori	362	134	496
2	Nu este încă hotărît	15	10	25
3	Nu dorește	48	33	81
Total		425	177	602

$$c = \sqrt{1 - \frac{1}{\frac{\frac{130044}{425} + \frac{17956}{177}}{496} + \frac{\frac{225}{425} + \frac{100}{177}}{25} + \frac{\frac{2304}{425} + \frac{1089}{177}}{81}}}$$

$$\approx \sqrt{1 - \frac{1}{1{,}01}} \approx 0{,}10.$$

Aşadar, coeficienţii de corelaţie care variază între limitele 0 şi 1 măsoară relaţiile corelaţionale între variabile, mai slabe sau mai tari, după cum se apropie de 0 sau de 1.

3.7. CORELAŢIA RANGURILOR

În cazul unor distribuţii care nu sînt de tip normal sau atunci cînd nu cunoaştem în prealabil parametrii fenomenelor cercetate vom utiliza anumite metode cunoscute sub denumirea de *metode neparametrice* [2]. Aceste metode elimină ipoteza privind tipul curbei de distribuţie şi ca urmare dau posibilitatea estimării celor mai variate feluri de distribuţii pe baza unor calcule destul de simple [11], [15].

Atunci cînd se utilizează metodele neparametrice nu se operează cu valorile caracteristicilor studiate, ci cu rangurile acestor valori, după ce în prealabil valorile caracteristicilor s-au trecut într-o ordine crescătoare sau descrescătoare. Deoarece se operează cu rangurile şi nu cu valorile iniţiale ale caracteristicilor, aceste metode permit măsurarea intensităţii relaţiilor statistice atît pentru caracteristicile cantitative, cît şi pentru caracteristicile calitative.

Calculul corelaţiei rangurilor se poate efectua cu ajutorul coeficienţilor lui S p e a r m a n şi lui K e n d a l l [11].

Formula de calcul a coeficientului de corelaţie al lui Spearman este dată în relaţia [11]:

$$R = 1 - 6 \frac{\Sigma d^2}{n(n^2 - 1)},$$

în care:

d — diferenţa dintre rangurile caracteristicilor x şi y;

n — numărul termenilor seriei.

---

[2] Metoda analizei neparametrice a fost elaborată iniţial de K. Pearson şi dezvoltată ulterior de Hotellong, Spearman, Kendall şi alţii.

Scanned with OKEN Scanner

Coeficientul corelației rangurilor se interpretează la fel ca și coeficientul corelației liniare, deoarece dacă notăm rangurile cu simbolurile utilizate pentru valorile caracteristicilor corelate, formula coeficientului lui Spearman se deduce din expresia:

$$R = \frac{\Sigma (x - \bar{x})(y - \bar{y})}{\sqrt{\Sigma(x - \bar{x})^2 (y - \bar{y})^2}}.$$

Această expresie este identică cu cea a coeficientului de corelație liniară (r). Cele două serii x și y, fiind constituite din primii n întregi, avem

$$\bar{x} = \bar{y} = \frac{n+1}{2}$$

și $$\Sigma(x - \bar{x})^2 = \Sigma(y - \bar{y})^2 = n\frac{n^2 - 1}{12},$$

iar $$(x - \bar{x})(y - \bar{y}) = \frac{1}{2}[\Sigma(x - \bar{x})^2 + \Sigma(y - \bar{y})^2] -$$

$$- \frac{1}{2}[\Sigma(x - \bar{x})(y - \bar{y})]^2 = n\frac{n^2 - 1}{12} - \frac{1}{2}\Sigma d^2,$$

în care:

$$d = (x - \bar{x}) - (y - \bar{y}) = x - y.$$

Coeficientul lui Spearman poate lua valori între -1 și $+1$ [11].

Spre exemplificare ne propunem să măsurăm relația dintre venitul național pe locuitor în 1965 (în dolari) în cîteva țări și născuții vii la 1000 de locuitori în diferite țări, utilizînd o serie de date statistice în acest scop[3]. Din analiza acestor date a reeșit concluzia că țările cu venit

[3] Exemplele sînt luate din:
— "Viața economică", nr. 13, 28 mar. 1968.
— Demographic Yearbook, United National, New York, 1967.

național scăzut pe un locuitor au un coeficient al natalității mai ridicat.

Datele de bază care au servit la calculul corelației rangurilor sînt cele înscrise în tabelul 14.

Tabelul 14

Țările	Venitul național pe un locuitor în 1965 (dolari)	Născuți la 1 000 locuitori în 1965	Rangul după :			
			Venitul național pe un locuitor în 1965 (R_x)	Născuții vii la 1000 locuitori (R_y)	Dif. $d_i = R_x - R_y$	d_i^2
S.U.A.	2893	18,5	1	12	−11	121
Canada	1825	19,6	2	10	− 8	64
Anglia	1451	17,9	3	13	−10	100
R.F.G.	1447	17,8	4	14	−10	100
Franța	1436	17,5	5	16	−11	121
Belgia	1406	15,8	6	17	−11	121
Austria	970	17,6	7	15	− 8	64
Italia	883	18,9	8	11	− 3	9
Argentina	740	23,0	9	8	1	1
Portugalia	351	22,3	10	9	1	1
Brazilia	217	43,0	11	6	5	25
Iran	211	48,0	12	2	10	100
Tunisia	179	45,1	13	4	9	81
Maroc	174	46,1	14	3	11	121
India	86	38,4	15	7	8	64
Haiti	80	45,0	16	5	11	121
Birmania	56	50,0	17	1	16	256
Total					Σd_i^2	1460

Înlocuind în formula coeficientului lui Spearman rezultatele obținute în tabelul 14 vom avea :

$$R = 1 - \frac{6 \Sigma d_i^2}{n(n^2 - 1)} = 1 - \frac{6 \times 1460}{17(17^2 - 1)} =$$

$$= \frac{8760}{4913 - 17} = -0{,}791.$$

Scanned with OKEN Scanner

Acest coeficient de corelaţie al rangurilor ne indică o relaţie statistică destul de intensă în sens invers (deoarece are semnul negativ) dintre cele două caracteristici studiate ; în timp ce valoarea caracteristicii x scade, valoarea caracteristicii y creşte.

Cel de-al doilea coeficient de corelaţie al rangurilor îl constituie coeficientul lui Kendall. Acest coeficient se calculează ca un raport dintre „scorul" total efectiv studiat, care reprezintă suma algebrică a numărului de ranguri superioare, respectiv inferioare, fiecărui rang luat succesiv al valorilor lui y, ordonate după seria valorilor lui x în ordine crescătoare sau descrescătoare şi „scorul" total maxim posibil care este egal cu $\frac{n(n-1)}{2}$, adică cu numărul combinărilor posibile din n elemente luate două cîte două.

Formula de calcul al acestui coeficient este următoarea :

$$\tau = \frac{2S}{n(n-1)} = \frac{S}{\frac{1}{2}n(n-1)},$$

în care :

S — suma algebrică a rangurilor superioare (P) şi inferioare (Q) care urmează fiecărui termen cuprins în coloana respectivă ;

n — numărul termenilor seriei.

Spre exemplificare, vom aplica coeficientul lui Kendall la datele calculate în tabelul 15.

Cifrele din coloana 7 s-au calculat efectuînd suma algebrică dintre P(col. 5) şi Q (col. 6), adică prin însumarea numărului rangurilor superioare cu cel al rangurilor inferioare ce urmează fiecărui termen cuprins în coloana respectivă.

Scanned with OKEN Scanner

Tabelul 15

Țările	Venit național pe un locuitor în 1965 (dolari)	Născuți la 1 000 locuitori în 1965	Rangul după		P	Q	S
			Venitul național pe un locuitor în 1965	Născuții vii la 1 000 locuitori			
0	1	2	3	4	5	6	7
S.U.A	2893	18,5	1	12	5	11	— 6
Canada	1825	19,5	2	10	6	9	— 3
Anglia	1451	17,9	3	13	4	10	— 6
R.F.G.	1447	17,8	4	14	3	10	— 7
Franța	1436	17,5	5	16	1	11	—10
Belgia	1406	15,8	6	17	—	11	—11
Austria	970	17,6	7	15	—	10	—10
Italia	883	18,9	8	11	—	9	— 9
Argentina	740	23,0	9	8	1	7	— 6
Portugalia	351	22,3	10	9	—	7	— 7
Brazilia	217	43,0	11	6	—	5	— 5
Iran	211	48,0	12	2	—	1	— 1
Tunisia	179	45,1	13	4	2	2	— 0
Maroc	174	46,1	14	3	—	1	— 1
India	86	38,4	15	7	—	2	— 2
Haiti	80	45,0	16	5	—	1	— 1
Birmania	56	50,0	17	1	—	—	— 1

Înlocuind valorile de mai sus în formula coeficientului lui Kendall (τ) vom avea [4]:

$$\tau = \frac{S}{\frac{1}{2}n(n-1)} = \frac{-85}{\frac{1}{2}17(17-1)} = \frac{-85}{\frac{17}{2}\cdot 16} = -0{,}625.$$

Acest coeficient (avînd semnul negativ) indică tot o relație statistică inversă între cele două caracteristici studiate, cu o intensitate mai slabă decît a coeficientului lui Spearman (—0,791).

Scanned with OKEN Scanner

În cazul cînd n este mare, între coeficientul lui Spearman şi cel al lui Kendall, există relaţia aproximativă [16]:

$$R \simeq \frac{3}{2}\tau.$$

De obicei, coeficientul lui Kendall indică o corelaţie mai moderată a caracteristicilor studiate, faţă de coeficientul lui Spearman, care are o valoare mai apropiată de 1 sau de -1.

În general, se preferă coeficientul lui Kendall care satisface într-o măsură mai mare cerinţele distribuţiei de selecţie pentru o valoare curentă a lui n şi pentru care este posibilă verificarea semnificaţiei [11].

3 8. ASOCIEREA CARACTERISTICILOR CALITATIVE

Informaţiile culese cu ajutorul unor metode care implică efectuarea unor observaţii măsurabile sînt susceptibile de o clasificare numerică definită; iar corespondenţa de intensitate a două caracteristici în acest caz, este stabilită prin rangul lor [17].

Realizarea unei ordonări de N elemente este posibilă prin:

a) punerea în evidenţă a prezenţei sau absenţei unei asocieri între caracteristica A şi B, avînd fiecare două caractere „da" şi „nu" (tabelul cu 2 linii şi 2 coloane, numit tabelul 2×2);

b) observarea gradului de asociere a două caracteristici avînd fiecare mai mult de două clase, cu ajutorul unui tabel general de contingenţă de m linii şi n coloane.

Scanned with OKEN Scanner

Tabelul 2×2 permite compararea între procentaje. De exemplu, să verificăm dacă există sau nu o asociere între sensibilitatea de intoxicare profesională a muncitorilor care lucrează într-o uzină chimică dintr-un oraș oarecare și locul unde aceștia își au domiciliul (în oraș, la țară) [4].

Schema unei astfel de asocieri s-ar putea pune în evidență cu ajutorul următoarei tabele 2×2 (elemente reale):

a	b	n_1
c	d	n_2
m_1	m_2	N

Pe baza elementelor din tabela 2×2 putem calcula abaterea-tip cu ajutorul formulei [4]:

$$\sigma_p = \sqrt{\frac{m_1}{N} \times \frac{m_2}{N}\left(\frac{1}{n_1} - \frac{1}{N}\right)} = \sqrt{\frac{m_1 m_2 n_2}{N^3 n_1}} =$$

$$= \sqrt{pq\left(\frac{1}{n_1} - \frac{1}{N}\right)}.$$

Problema constă în compararea a două procentaje, făcînd diferența între valoarea teoretică (p) și cea reală (p') a probabilităților:

$$p - p' = \frac{m_1}{N} - \frac{a}{n_1} = \frac{m_1 n_1 - aN}{N n_1}$$

Scanned with OKEN Scanner

Abaterea redusă, determinată ca raport între abaterea variabilei față de medie și abaterea tip, va fi :

$$t = \frac{p - p'}{\sigma p} = \frac{m_1 n_1 - na}{N n_1} : \sqrt{\frac{m_1 m_2 n_2}{N^3 n_1}} =$$

$$= \frac{m_1 n_1 - Na}{\sqrt{\frac{m_1 m_2 n_1 n_2}{N}}}, \text{ sau :}$$

$$t = \frac{(a + c)\ (a + b) - a(a + b + c + d)}{\sqrt{\frac{m_1 m_2 n_1 n_2}{N}}} =$$

$$= \frac{(bc - ad)\sqrt{N}}{\sqrt{m_1 m_2 n_1 n_2}}.$$

Cunoscînd valoarea lui t se poate determina funcția de repartiție normală Ø (t) utilizînd în acest scop tabela din anexa 1; în funcție de Ø (t) poate fi determinată densitatea de probabilitate [4] :

$$P = 1 - 2\,Ø\,(t).$$

În ipoteza de independență, numerele teoretice din căsuțele unei tabele de contingență pot fi determinate cu ajutorul formulelor :

$$\alpha = \frac{m_1 n_1}{N}\ ;\ \beta = \frac{m_2 n_1}{N}\ ;\ \gamma = \frac{m_1 n_2}{N}\ ;\ \delta = \frac{m_2 n_2}{N}\ ;$$

Scanned with OKEN Scanner

Introducem aceste numere în căsuțele tabelei de contingență (elemente teoretice):

α	β	n_1
γ	δ	n_2
m_1	m_2	N

Valoarea testului χ^2 va fi [4]:

$$\chi^2 = \frac{(a-\alpha)^2}{\alpha} + \frac{(b-\beta)^2}{\beta} + \frac{(c-\gamma)^2}{\gamma} + \frac{(d-\delta)^2}{\delta}.$$

În ipoteza că:

$$a - \alpha = \gamma - c = \beta - b = d - \delta,$$

expresia lui χ^2 devine:

$$\chi^2 = \frac{(ad-bc)^2 N}{n_1 n_2 m_1 m_2} = \frac{(ad-bc)^2(a+b+c+d)}{(a+b)(a+c)(b+c)(c+d)}.$$

Rezultă că valoarea testului χ^2 poate fi obținută fără a calcula coeficienții α, β, γ, δ.

Distribuția teoretică a parametrului χ^2 se poate calcula cu ajutorul tabelului din anexa 2 pentru un grad de libertate $\nu = 1$. Probabilitatea de a obține un dezacord este ca valoarea lui P să fie superioară cifrei 0,05. Diferența obținută se verifică dacă este sau nu semnificativă [4].

Scanned with OKEN Scanner

În cazul unui tabel de contingenţă $m \times n$, avînd m linii şi n coloane, fiecare căsuţă conţine un număr n_{rs} de indivizi care posedă simultan atît caracterul X_r cît şi caracterul Y_s.

Forma acestui tabel este analoagă, formei tabelului de corelaţie, cu o singură deosebire: caracterele nu sînt exprimate printr-o variabilă numerică [4].

CARACTERE X

CARACTERE Y	x_1	x_2	x_3	...	x_r	...	x_m	Total
y_1	n_{11}	n_{21}	n_{31}	...	n_{r1}	...	n_{m1}	n_1
y_2	n_{12}	n_{22}	n_{32}	...	n_{r2}	...	n_{m2}	n_2
y_3								
...								
y_s					n_{rs}			ns
...								
y_n								
Total					n_r			N

În cazul independenţei complete între caractere se poate scrie relaţia [4]:

$$n_{rs} = n_r \frac{n_s}{N}$$

În general, există o diferență între observație și ipoteză, adică:

$$\Sigma\left(n_{rs} - \frac{n_r n_s}{N}\right) \neq 0.$$

În cazul general, cînd admitem ipoteza de independență, efectul teoretic al unei căsuțe $X_r Y_s$ va fi:

$$n'_{rs} = \frac{n_r . n_s}{N}$$

Totalul diferențelor divergente $(n' - n)$ exprimă un dezacord global între observație și ipoteză [18].

Asociația dintre două caractere va fi dată de cantitatea $\chi^2 = \Sigma \frac{(n' - n)}{n'}$, care este nulă în ipoteza de independență, iar legea de distribuție a lui χ^2 este pentru n': $\nu = (m - 1)(n - 1)$ grade de libertate.

3.9. CORELAȚIA PARȚIALĂ

Pentru analiza științifică a corelației prezintă interes și determinarea gradului de asociere dintre variabila dependentă (x) și o variabilă independentă (y), cu excluderea influenței celorlalte variabile independente (z_i). Asemenea legături între variabile se măsoară de obicei cu ajutorul *coeficientului de corelație parțială*. Formula de calcul a acestui coeficient este următoa-

Scanned with OKEN Scanner

rea [11]:

$$r_{xy.z} = \frac{r_{xy} - r_{xy}\, r_{yz}}{\sqrt{1 - r_{xz}^2}\,\sqrt{1 - r_{yz}^2}}$$

Această formulă se poate aplica atunci cînd studiem o corelaţie parţială între variabilele x şi y pentru z invariabil.

Dacă însă vrem să calculăm corelaţia parţială între două variabile, considerînd efectul celorlalte variabile care le influenţează ca fiind constant, estimarea coeficientului cu ajutorul unei formule numerice este deosebit de complicată.

Atunci cînd se aplică coeficientul de corelaţie parţială, eliminîndu-se m variabile din analiză, putem determina eroarea-standard cu ajutorul formulei [4]:

$$\sigma_z = \frac{1}{\sqrt{n - k - 1}},$$

în care:

n — numărul cuplurilor observate;
k — numărul total de variabile.

Efectul unui număr oarecare de alte variabile z_1, z_2, z_3 etc. se poate elimina simultan din valorile lui x şi y, obţinîndu-se coeficientul de corelaţie parţială $r_{xy.\,z_1.\,z_2.\,z_3.\,\ldots.\,z_m}$.

De obicei calculul se simplifică dacă se utilizează determinantul [4]:

$$R = \begin{vmatrix} 1 & r_{12} & r_{13} \ldots r_{1k} \\ r_{21} & 1 & r_{23} \ldots r_{2k} \\ \ldots\ldots\ldots\ldots \\ r_{k1} & r_{k2} & r_{k3} \ldots 1 \end{vmatrix}$$

Scanned with OKEN Scanner

Elementele determinantului R sînt coeficienții de corelație simplă dintre variabile.

Coeficientul de corelație parțială între două variabile, cu excluderea celorlalte variabile independente, se poate calcula cu ajutorul formulei [4]:

$$r_{xy,\, z_1 z_2, \ldots, z_m} = \frac{R_{xy}}{\sqrt{R_{xx} \cdot R_{yy}}},$$

unde:

R_{xy} este minorul corespunzător lui r_{xy} din determinantul R

3.10. CORELAȚIA MULTIPLĂ

Legăturile dintre variabilele unei structuri sociale se caracterizează prin faptul că variabilele variază sub influența unui complex de factori, dintre care unii sînt esențiali iar alții neesențiali. Determinînd efectul factorilor esențiali care influențează un proces social oarecare se pot lua decizii pentru asigurarea condițiilor optime de dezvoltare a acestuia, intensificînd influența factorilor pozitivi și diminuînd sau eliminînd influența factorilor negativi. Calculele de punere în evidență a corelației multiple servesc la aprecierea rolului unui ansamblu de factori în evoluția proceselor sociale deosebit de complexe [19], [20].

Influența comună a variabilelor unei structuri sociale se poate măsura cu ajutorul indicatorilor corelației multiple. Un astfel de indicator este *coeficientul corelației multiple*, cu ajutorul căruia se poate măsura influența simultană a variabilelor independente $x_1, x_2, \ldots, x_n$ asupra variabilei dependente y. Formula de calcul este [11]:

$$R_{y,\, x_1, \ldots, x_n} = \sqrt{1 - \frac{\Sigma (y - \overline{y}_{x_1, \ldots, x_n})^2}{\Sigma (y - \overline{y})^2}},$$

Scanned with OKEN Scanner

în care :

y — variabila dependentă ;
$\bar{y}$ — media aritmetică a variabilei dependente ;
$\bar{y}_{x_1, \ldots, x_n}$ — media aritmetică a variabilei dependente exprimată printr-o funcție liniară de tipul $\bar{y}_{x_1 \ldots x_n} = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_n x_n$, unde a_0 și a_j $(j = 1,2, \ldots n)$ sînt parametrii liniei de regresie liniară multiplă.

Valoarea coeficientului de corelație multiplă depinde de raportul dintre dispersia valorilor calculate pe baza ecuației de regresie multiplă și dispersia valorilor observate ale variabilei dependente. Cu cît valorile observate se abat mai puțin de la linia de regresie multiplă, cu atît coeficientul de corelație are o valoare mai mare și deci corelația este mai intensă. Dacă admitem ipoteza că $\Sigma (y - \bar{y}_{x_1, x_2, \ldots, x_n})^2 = 0$, atunci coeficientul de corelație $R_{y, x_1 x_2, \ldots, x_n} = 1$, adică valorile observate ale lui y, coincid cu valorile calculate pe baza ecuației de regresie multiplă. Datorită acestui fapt, variația variabilei y este integral explicată prin variația variabilelor independente în calculul corelației multiple.

Coeficientul corelației multiple poate fi dedus și din coeficienții corelației simple, dacă se consideră [11] :

$$\sigma^2_{y.x_1 x_2 \ldots x_n} = \sigma^2_y (1 - r^2_{yx_1}) (1 - r^2_{y x_2 . x_1}) (1 - r^2_{y x_3 . x_1 x_2}) \ldots$$

$$\ldots (1 - r^2_{y x_n . x_1 x_2 \ldots x_{n-1}}).$$

Pentru cazul a trei variabile vom avea :

$$\frac{\sigma^2_{y x_2 x_3}}{\sigma^2_y} = (1 - r^2_{y x_2}) (1 - r^2_{y x_3 x_2}) =$$

$$= \frac{1 - r^2_{y x_2} - r^2_{x_2 x_3} - r^2_{y x_3} + 2 r_{y x_2} r_{x_2 x_3} r_{y x_3}}{1 - r^2_{x_2 x_3}}.$$

De unde rezultă :

$$R_{y\,x_2\,x_3} = \sqrt{\frac{r^2_{y\,x_2} + r^2_{y\,x_3} - 2r_{y\,x_2} \cdot r_{y\,x_3}\, r_{x_2\,x_3}}{1 - r^2_{x_2\,x_3}}}$$

Față de coeficienții de corelație simplă, coeficientul de corelație multiplă prezintă un mare interes, deoarece el exprimă influența totală și concomitentă a variației unui grup de variabile (indicatori).

3.11. DETERMINAȚIA MULTIPLĂ

Metodele statistico-matematice permit determinarea contribuției fiecărui factor în parte la modificarea fenomenului-efect. Indicatorii utilizați în acest scop poartă denumirea de *coeficienți de determinație* [11].

Gradul de influență a fiecărui factor inclus în analiza corelațională se exprimă prin acea parte a dispersiei valorilor fenomenului efect care este determinată de variația valorilor factorului respectiv. Cu alte cuvinte, descompunerea influenței pe factori constă în descompunerea dispersiei variabilei dependente σ^2_y pe seama dispersiei σ^2_x a fiecărei variabile independente. În ipoteza determinării tuturor factorilor care influențează fenomenul cercetat, dispersia variabilei dependente σ^2_y se descompune integral pe factori.

Ca unitate de măsură a influenței simultane, exercitată de variația tuturor factorilor analizați, se utilizează pătratul coeficientului de corelație multiplă. Această mărime este cunoscută sub denumirea de *coeficient al determinației multiple* care pentru trei variabile poate fi calculat prin relația [11] :

$$r^2_{y.x_1\,x_2} = 1 - \frac{\sigma^2_{y.x_1\,x_2}}{\sigma^2_y},$$

în care :

$r^2_{y \cdot x_1 x_2}$ reprezintă partea dispersiei σ^2_y explicată prin influenţa variabilelor x_1 x_2, adică mărimea cu care σ^2_y se micşorează datorită influenţei comune a celor două variabile x_1 şi x_2.

Notînd cu $d_{y \cdot x_1 \ldots x_n}$ coeficientul de determinaţie multiplă, putem scrie :

$$d_{y.x_1 x_2 \ldots x_n} = r^2_{y.x_1 x_2 \ldots x_n} = 1 - \frac{\sigma^2_y - \sigma^2_{y.x_1 x_2 \ldots x_n}}{\sigma^2_y}$$

Coeficientul $d_{y\, x_1 x_2 \ldots x_n}$ poate varia în intervalul 0 şi 1, arătînd cît din dispersia indicatorului y este determinată de influenţa simultană a variaţiei celorlalţi indicatori $x_1 x_2 \ldots x_n$.

Influenţa separată a unui factor $x_i (i=1, 2, \ldots, n)$ se poate determina precum urmează :

$$d_{y \cdot x_i} = 1 - \frac{\sigma^2_y - b\left(\frac{\sum y x_i}{n} - \bar{y}\bar{x}_i\right)}{\sigma^2_y} = b \cdot \frac{\frac{\sum y x_i}{n} - \bar{y}\,\bar{x}_i}{\sigma^2_y},$$

unde :

b — coeficientul liniei de regresie ;

$\bar{x}_i$; $\bar{y}$ — valorile medii ale variabilelor x_i şi ale variabilei y.

Suma coeficienţilor $d_{y\, x_i}$ este egală cu determinaţia totală atribuită acţiunii simultane a tuturor variabilelor (x_i) care exercită un efect asupra variabilei determinate (y).

Scanned with OKEN Scanner

3.12. VERIFICAREA SEMNIFICAȚIEI COEFICIENTULUI DE CORELAȚIE MULTIPLĂ

Calculul corelației multiple se efectuează adeseori pe baza informațiilor obținute dintr-o cercetare selectivă, astfel încît indicatorii obținuți nu constituie decît estimații ale valorii lor exacte în colectivitatea generală. Prin urmare, este absolut necesar să se cunoască precizia cu care coeficientul corelației multiple a fost estimat, verificîndu-se semnificația lui și determinîndu-se limitele de încredere corespunzătoare unui anumit nivel de semnificație [11].

Coeficientul de corelație multiplă fiind o mărime normată, funcția de frecvență $f(r)$ a unei variabile oarecare, depinde numai de volumul selecției (numărul perechilor de observații, n) și de valoarea coeficientului de corelație teoretic (ρ) [11], [21].

În selecțiile din colectivitatea inițială normală, repartițiile se apropie de cea normală, pe măsură ce crește n. Această tendință apare mai clar pentru valorile ρ apropiate de zero ale coeficientului de corelație al colectivității generale, decît pentru ρ diferit mult de zero. Pentru valorile lui ρ apropiate de $+1$ sau de -1, este necesar ca n să fie destul de mare pentru ca distribuția lui r să fie simetrică și aproximativ normală. O dată cu apropierea de zero a coeficientului de corelație și cu creșterea numărului perechilor de observații, coeficientul de corelație empiric are o repartiție aproximativ normală cu $M(r) = \rho$ și dispersia $\sigma_r = \frac{1 - \rho^2}{\sqrt{n - 1}}$, unde σ_r este eroarea pătratică a lui r. Cînd verificăm ipoteza că selecția dată este luată dintr-o colectivitate generală în care ρ este nul, formula se reduce la $\sigma_r = \frac{1}{\sqrt{n - 1}}$. Dacă între coeficientul de corelație calculat (R) și abaterea pătratică (σ_r) există o diferență mare, ipoteza de nulitate a lui ρ nu poate fi admisă.

Scanned with OKEN Scanner

În continuare, se poate determina eroarea-limită Δ_r şi limitele de încredere pentru coeficientul de corelaţie în colectivitatea generală, considerînd r aproximativ egal cu ρ, după cum urmează [4]:

$$\Delta r = \frac{1 - R^2}{\sqrt{n - 1}}$$

Limitele de încredere pentru coeficientul de corelaţie în colectivitatea generală vor fi:

$$R \pm t\Delta_r,$$

unde criteriul t reprezintă raportul dintre coeficientul de regresie şi eroarea de selecţie (se adoptă în funcţie de probabilitatea P cu care se garantează apariţia erorilor).

Semnificaţia coeficientului de corelaţie multiplă poate fi verificată şi cu ajutorul testului F al lui Snédécor [22].

$$F = \frac{n - k}{k - 1} \cdot \frac{R^2}{1 - R^2},$$

cu $n_1 = k - 1$ şi $n_2 = n - k$ — grade de libertate.

Testul de semnificaţie a coeficientului corelaţiei multiple constă în a verifica dacă variaţia datorită legăturii este sau nu superioară variaţiei reziduale, dacă F este sau nu semnificativ superior faţă de unitate, cu considerarea gradelor de libertate respective [4].

Coeficientul de corelaţie multiplă (R) este semnificativ dacă variaţia explicată prin regresie (variaţia dintre grupe) este semnificativ superioară variaţiei reziduale (variaţiei în interiorul grupelor) [23].

3.13. CONCLUZII

Constatarea unui simplu paralelism în manifestarea şi dezvoltarea de masă a unor fenomene sociale nu trebuie neapărat să ne conducă la concluzia că între aceste feno-

Scanned with OKEN Scanner

mene ar exista o relație de cauzalitate. Sînt frecvente cazurile cînd, deși există o cale comună de dezvoltare a două sau mai multe fenomene sociale, totuși între ele să nu existe o relație de cauzalitate.

În majoritatea cazurilor, relațiile de dependență între caracteristicile unui proces social oarecare nu apar într-o formă simplă, pură ; în marea lor majoritate, aceste relații sînt multiple și complexe, dificil de identificat. De aceea, în decelarea relațiilor, sociologul trebuie să se folosească de metoda abstracției științifice, care-i permite ca, în cursul cercetării, să izoleze cazurile neesențiale ce ar putea să influențeze asupra rezultatelor obținute și să se oprească asupra factorilor esențiali.

Este cunoscut că structurile sociale sînt deosebit de complexe, fiind determinate de un ansamblu de caracteristici și de relații statistice. De aceea, studiul izolat al unor relații statistice între caracteristicile unei structuri sociale date, rupte de sistemul de relații din care fac parte, nu poate procura informații sigure despre intensitatea acestor relații.

Pentru realizarea unei cunoașteri complete a unei structuri sociale date este necesar să se determine gradul de tărie al relațiilor statistice între caracteristicile care determină o astfel de structură deosebit de complexă. Relațiile statistice pot fi privite drept corelații statistice, dacă avem în vedere asocierea existentă între caracteristicile unei structuri sociale.

Într-o cercetare sociologică uneori sîntem nevoiți nu numai să luăm în considerare relațiile statistice multilaterale dintre caracteristicile unei structuri sociale, ci să „rupem" într-o anumită măsură această structură în scopul de a studia unele corelații puternice, mai importante. În cadrul unei astfel de cercetări sîntem nevoiți să ținem cont nu numai de existența pură a corelațiilor statistice, ci și de faptul că ele sînt esențiale sau neesențiale pentru structura socială studiată.

Scanned with OKEN Scanner

Este inadmisibil ca într-o cercetare sociologică să facem abstracţie de existenţa unor relaţii statistice necesare. Dar dacă nu se face abstracţie de relaţiile statistice întîmplătoare pot fi comise erori foarte mari, care pot conduce la concluzii eronate într-o astfel de cercetare.

În domeniul cercetărilor sociologice, abstractizarea ştiinţifică se realizează într-o măsură însemnată prin utilizarea metodelor statistico-matematice. Această fază a „abstractizării" va trebui în final să fie depăşită prin trecerea la „concret". În procesul cercetării ştiinţifice această trecere înseamnă, de fapt, separarea factorilor după importanţa influenţei lor, făcîndu-se abstracţie de toţi ceilalţi factori, în afară de cei de importanţă deosebită.

Cercetarea corelaţiilor simple este indispensabilă sociologului în faza de analiză statistico-matematică a unui proces social oarecare. Rezultatele obţinute însă trebuie interpretate ca o variaţie concomitentă a două variabile, fără a avea pretenţia ca prin aceasta să cunoaştem care dintre ele este cauza şi care efectul. Dacă totuşi vrem să aflăm acest lucru este necesar să apelăm la modelele matematice complexe, de mare rafinament.

Prin metoda corelaţiei multiple se înţelege întreaga metodologie de cercetare a interdependenţelor şi conexiunilor în care se găsesc fenomenele sociale.

În analiza corelaţională multiplă legăturile dintre mai multe variabile sînt exprimate sub formă de indicatori sintetici, cu ajutorul cărora sînt caracterizate numeric intensităţile acestor legături.

Relaţiile dintre variabile reflectă de fapt legăturile cu caracter complex, existente în mod obiectiv între fenomenele de masă. Caracterul complex al dependenţelor statistice între variabile pune pe primul plan problema identificării corelaţiilor multiple dintre fenomene. Calculul indicatorilor corelaţionali este admis cu condiţia stabilirii anticipate a legăturilor cauzale între fenomenele cercetate.

Scanned with OKEN Scanner

În același timp, un simplu paralelism în variația mai multor fenomene nu înseamnă că între acestea există relații cauzale.

În scopul obținerii unor concluzii suficient de întemeiate, pe baza calculelor de corelație, este necesară includerea în cercetare, pe cît posibil, a tuturor factorilor cu acțiune sistematică, esențială. În măsura în care această condiție nu este satisfăcută, valoarea de cercetare a indicatorilor de corelație multiplă este diminuată într-o măsură importantă.

BIBLIOGRAFIE

1. Prof. Dr. C. Ionescu ș. a., *Metode statistice de cercetare a corelațiilor în economie*, Edit. Academiei, 1968.
2. Gh. Mihoc și V. Urseanu, *Matematici aplicate în statistică*, Edit. Academiei, 1962.
3. C. A. Moser, *Metodele de anchetă în investigarea fenomenelor sociale*, Edit. științifică, 1967.
4. P. Pèpe și M. Tisserand-Perrier, *Méthodes statistiques dans les sciences sociaux*, Paris, Masson & Cie, 1962.
5. A. G. Worthing și J. Jeffner, *Prelucrarea datelor experimentale* (traducere).
6. A. Lorfon, *Initiation pratique à la statistique*, Paris, Eyrolle et Gauthier-Villars, 1956.
7. F. Egermayer, *Metoda de analiză a corelației în cercetările economice*, în : Rev. statistică, nr. 7, 1961.
8. L. Tövissi și H. Rancu, *Elemente de teorie a corelației*, în : Revista de statistică, nr. 5—6, 1960.
9. A. Vessereau, *La statistique*, Paris, Presses Universitaires, Collection „Que sais-je?", 1945.
10. A. De Moncetz, *Initiation aux méthodes de la statistique*, Société d'édition française et internationales, 1945.
11. I. Marinescu ș. a., *Elemente de statistică matematică și aplicațiile ei*, București, Edit. științifică, 1966.
12. Manea Mănescu și M. Vasilescu (sub. red.). *Statistica nivelului de trai al populației*. București, Edit. științifică, 1969.
13. P. Pèpe, *Présentation des statistiques*, Paris, Dunod, 1959.

Scanned with OKEN Scanner

14. L. B. Itelson, *Metode matematice și cibernetice în pedagogie.* București, Edit. didactică și pedagogică, 1967.
15. R. A. Fisher, *Les méthodes statistiques adoptées à la recherche scientifique*, Paris, Presses Universitaires, 1947.
16. G. Yule și Kendall, *An introduction to the theory of statistics*, London, 1950.
17. P. Pèpe, *Procédés et matériels de dépouillements statistiques*, Paris, Dunod, 1961.
18. K. Szaniawski, *Metody statyczne w socjologii*, Warszawa, Panstwowe nijdawnictwo naukiwe, 1968.
19. V. Trebici ș. a., *Aplicații ale metodei corelației în analiza statistică a productivității muncii*, în: Probleme economice, nr. 4, 1962.
20. G. Darmois, *L'analyse de corrélations*, Paris, Conférences du Palais de la Decouverte, 1948.
21. A. Monjallon, *Introduction à la méthode statistique*, Paris, Vuibert, 1954.
22. J. Dobos, *Méthodologie économique et technique statistique*, IV, Montpellier, 1965.
23. P. P. Maslov, *Sociologija i statistika.* Moskva, 1967, „Statistika" 8.

Scanned with OKEN Scanner

ANEXA 1

Funcțiile de repartiție normale (Ø (t)) și densitățile de probabilitate (1—2 Ø (t))

t	Ø(t)	1—2 Ø (t)	t	Ø(t)	1—2 Ø (t)
0,00	0,	0,	1,45	0,4265	0,1470
0,05	0,0199	0,9602	1,50	0,4332	0,1336
0,10	0,0398	0,9204	1,55	0,4394	0,1212
0,15	0,0596	0,8808	1,60	0,4452	0,1096
0,20	0,0793	0,8414	1,65	0,4505	0,0990
0,25	0,0987	0,8026	1,70	0,4554	0,0892
0,30	0,1179	0,7642	1,75	0,4599	0,0802
0,35	0,1368	0,7264	1,80	0 4641	0,0718
0,40	0,1554	0,6892	1,85	0,4678	0,0644
0,45	0,1736	0,6528	1,90	0,4713	0,0574
0,50	0,1915	0,6170	1,95	0,4744	0,0512
0,55	0,2088	0,5824	2,00	0,4772	0,0456
0,65	0,2422	0,5156	2,10	0,4821	0,0358
0,70	0,2580	0,4840	2,20	0,4861	0,0278
0,75	0,2734	0,4532	2,30	0,4893	0,0214
0,80	0,2881	0,4238	2,40	0,4918	0,0164
0,85	0,3023	0,3954	2,50	0,4938	0,0124
0,90	0,3159	0,3682	2,60	0,4953	0,0094
0,95	0,3289	0,3422	2,70	0,4965	0,0070
1,00	0,3413	0,3174	2,80	0,4974	0,0052
1,05	0,3531	0,2938	2,90	0,4981	0,0038
1,10	0,3643	0,2714	3,00	0,49865	0,0027
1,15	0,3749	0,2502	3,20	0,49931	0,00138
1,20	0,3849	0,2302	3,40	0,49966	0,00068
1,25	0 3944	0,212	3,60	0,499841	0,000318
1,30	0,4032	0,1936	3,80	0,499928	0,000144
1,35	0,4115	0,1770	4,00	0,499968	0 000064
1,40	0,4192	0,1616	4,50	0,499997	0,000006

Scanned with OKEN Scanner

ANEXA 2

Testul χ^2 a lui Pearson

Gradul de libertate	P = 0,05 χ^2 (95)	P = 0,01 χ^2 (99)	Gradul de libertate	P = 0,05 χ^2 (95)	P = 0,01 χ^2 (99)
1	3,84	6,64	16	26,30	32,00
2	5,99	9,21	17	27,60	33,41
3	7,82	11,34	18	28,87	34,80
4	9,49	13,28	19	30,14	36,19
5	10,07	15,09	20	31,41	37,57
6	12,59	16,81	21	32,67	38,93
7	14,07	18,48	22	33,92	40,29
8	15,51	20,09	23	35,17	41,64
9	10,92	21,67	24	36,42	42,98
10	18,31	23,21	25	37,65	44,31
11	19,68	24,72	26	38,89	45,64
12	21,03	26,22	27	40,11	46,96
13	22,36	27,69	28	41,34	48,28
14	23,69	29,14	29	42,56	49,59
15	25,00	30,58	30	43,77	50,89

Scanned with OKEN Scanner

4. MODELE DE ANALIZĂ CAUZALĂ

4.1. INTRODUCERE

În cercetarea problemelor de inferenţă între variabilele care descriu fenomenele şi procesele sociale se apelează frecvent la noţiunile de corelaţie statistică şi de relaţie cauzală, cunoscîndu-se faptul că ultima noţiune se poate deduce din prima. Cu toate acestea, între aceste două noţiuni există o deosebire evidentă, deoarece prin corelaţie statistică vom înţelege o variaţie concomitentă a două sau mai multe variabile, printr-o relaţie cauzală vom pune în evidenţă o dependenţă dintre variabile, din care cel puţin una este cauza unui efect exercitat asupra celeilalte. Deoarece corelaţiile statistice dintre variabile constituie proba relaţiilor cauzale, discuţiile asupra structurilor cauzale sînt oportune doar în prezenţa unor corelaţii statistice demonstrate.

Participarea corelaţiilor statistice la realizarea inferenţelor cauzale a determinat pe numeroşi autori, printre care Simon, Blalock, Coleman şi Boudon, să utilizeze corelaţiile statistice la deducerea relaţiilor cauzale [1]. În acest scop, aceşti autori au apelat la analiza variaţiilor concomitente în studiul structurilor cauzale. Prin efectuarea acestui studiu, au demonstrat că deducerea relaţiilor cauzale din corelaţiile statistice este pe deplin posibilă. Descoperirea unei interdependenţe între corelaţiile statistice şi relaţiile cauzale a deschis perspective excepţionale în studiul structurilor cauzale cu ajutorul unor *modele matematice de inferenţe cauzale dintre variabile* (modele cauzale), fundamentate pe metode statistico-matematice perfecţionate.

Scanned with OKEN Scanner

Fenomenele şi procesele sociale sînt considerate sisteme de relaţii cauzale deosebit de complexe, sisteme care pot fi formulate matematic. Formularea matematică a relaţiilor cauzale nu constituie o formulare în limbajul curent, ci o formulare matematică aptă să ţină seama de proprietăţile formale ale cauzalităţii implicate într-o relaţie cauzală. Formularea matematică a structurilor cauzale prezintă dificultăţi serioase. Dintre acestea, reţinem pe cele mai importante [1]:

a) este extrem de dificilă formularea fidelă a relaţiilor cauzale, puse în evidenţă în grafele orientate, cu ajutorul unor ecuaţii algebrice, adică prin utilizarea unor relaţii simetrice prin excelenţă. De altfel, formularea matematică a cauzalităţii dată de R. Boudon a generat o serie de discuţii pe această temă între K. Polk, H. M. Blalock şi W. S. Robinson [1];

b) exprimarea matematică a relaţiei cauzale se face la R. Boudon cu ajutorul unui coeficient pe care-l denumeşte de „dependenţă", care reprezintă un caz particular al coeficientului de regresie liniară, atunci cînd structura cauzală studiată este identificabilă [2];

făcînd o distincţie evidentă între coeficientul de dependenţă şi coeficientul de regresie, R. Boudon a preferat pe cel dintîi;

c) dacă la prima vedere, existenţa unei relaţii de dependenţă între două variabile poate fi verificată, existenţa cauzei şi efectului într-o astfel de relaţie pare dificil de verificat [3].

Dificultăţile generale ale metodelor actuale de analiză a inferenţelor cauzale dintre variabile implică o reexaminare a problemelor menţionate, în scopul perfecţionării actualului mod de măsurare a gradului de dependenţă între variabile, astfel încît aceasta să permită descoperirea cauzei şi a efectului produs într-o relaţie cauzală.

Simplitatea unor structuri cauzale nu poate fi decît aparentă. De aceea, structurile cauzale simple, adică acelea în care variabilele explicative nu depind unele de altele

sau fac aceasta într-un mod parazitar, eventualele „perturbaţii" putîndu-se elimina sau controla, sînt extrem de rare în domeniul sociologiei.

În orice analiză a modelului cauzal trebuie să pornim de la principiul pluralităţii posibile a cauzelor, iar cazul cauzei unice nu poate fi decît un caz limită.

Relaţiile cauzale fundamentate de R. Boudon şi H.Simon au la origine o serie de proceduri econometrice, proceduri care au întotdeauna ca obiect o reţea complexă de cauze şi de efecte.

Într-o structură cauzală complexă, o expresie ca „x este cauza lui y" este cu totul nepotrivită. Într-adevăr, într-o astfel de structură vom considera cauza globală a lui y ca „ansamblul cauzal al lui y". În acest caz, relaţia existentă între un x particular (x_j) şi y poate fi măsurată prin efectuarea unor intervenţii experimentale în care se variază sistematic una din cauze, fără a afecta pe celelalte (disociabilitate artificială) sau prin punerea în evidenţă a interdependenţei variaţiilor cauzale între ele şi pe baza unor observaţii organizate în acest scop (disociabilitate naturală). În cazul structurilor cauzale complexe cel mai adecuat model cauzal este cel de tipul „x este doar una din cauzele lui y". Trecerea de la situaţia cauzalităţii la tratarea matematică nu se poate face însă fără a ţine seama de efectul celorlalte cauze din sistem. De regulă, trecerea de la inferenţă cauzală la modelul ei matematic se face prin intermediul *coeficientului de dependenţă*, prezentat de R. Boudon ca o extindere a coeficientului de regresie [4]. Forma diferită de calcul a coeficientului de dependenţă în comparaţie cu coeficientul de regresie îl determină pe R. Boudon să facă o distincţie netă între aceşti doi coeficienţi. Cu toate acestea, se întîlnesc situaţii în care coeficienţii de dependenţă sînt cazuri particulare ale coeficienţilor de regresie, ca de pildă, atunci cînd structura cauzală este identificabilă.

În studiul structurilor cauzale complexe ne interesează în mod deosebit să cunoaştem care este efectul pe care-l exercită fiecare variabilă din sistem asupra variabilei

dependente pe care o analizăm cauzal. În acest scop este necesar să elaborăm un model matematic al inferenţelor cauzale din structura socială analizată. Trecerea de la situaţia cauzalităţii la tratarea matematică se poate face prin intermediul coeficienţilor de dependenţă şi de influenţă cauzală. Aceşti coeficienţi permit verificarea diferitelor sisteme cauzale din structurile sociale analizate cu ajutorul metodelor matematice.

4.2. NOŢIUNILE DE CORELAŢIE STATISTICĂ ŞI DE RELAŢIE CAUZALĂ

În relaţiile dintre variabilele unei structuri sociale ne interesează atît gradul de tărie al corelaţiei cît şi sensul relaţiei cauzale. De aceea, cunoaşterea lor este absolut necesară în cercetarea relaţiilor existente între variabile, dacă avem în vedere faptul că între aceste două noţiuni există o deosebire evidentă [3].

La capitolul 3 prin corelaţie statistică s-a definit o variaţie concomitentă a două variabile, fără ca prin studiul acestei variaţii să elucidăm sensul legăturii de cauză-efect sau gradul de tărie al relaţiei cauzale dintre aceste variabile. De exemplu, între producţia globală industrială şi producţia agricolă există o corelaţie statistică, fără a exista o relaţie cauzală între aceşti indicatori. Corelaţia statistică scoasă în evidenţă se poate datora influenţei unei a treia variabile : nivelul de dotare cu fonduri fixe productive a industriei şi a agriculturii. Prin urmare, în acest caz una din cauzele care generează o corelaţie statistică între volumul producţiei industriale şi cel al producţiei agricole este nivelul diferit de dotare cu fonduri fixe productive a acestor ramuri ale producţiei materiale. De altfel, explicaţia reală ar fi că un volum sporit de fonduri fixe productive determină un spor de producţie în indus-

Scanned with OKEN Scanner

trie şi în agricultură. Datorită acestui fapt, producţia globală industrială este corelată cu producţia agricolă, fără a exista o relaţie cauzală între ele, deoarece fiecare este influenţată cauzal de către dotările cu fonduri fixe efectuate în acest scop.

Un alt exemplu îl constituie corelaţia existentă între nivelul de calificare al muncitorilor şi nivelul productivităţii muncii lor. Această corelaţie între cei doi indicatori, arată că, pe măsura creşterii calificării muncitorilor are loc creşterea productivităţii muncii. Relaţia existentă, însă, nu se datorează în exclusivitate efectului ce-l exercită creşterea calificării muncitorilor asupra creşterii productivităţii muncii. Desigur că creşterea productivităţii muncii se poate datora şi altor factori ca : reducerea fluctuaţiei muncitorilor, creşterea gradului de înzestrare tehnică, sporirea indicelui de utilizare al maşinilor etc. În virtutea acestui fapt, corelaţia statistică dintre nivelul de calificare al muncitorilor şi nivelul productivităţii muncii nu reflectă o dependenţă totală între aceşti doi indicatori, deoarece productivitatea muncii este un indicator sintetic ; nivelul acestui indicator este rezultatul efectului ce-l exercită un ansamblu de factori de natură diferită : tehnici, tehnologici (grad de automatizare, de mecanizare), economici (salariul mediu, fluctuaţia etc.), sociali (vîrsta muncitorului, vechimea în cîmpul muncii, vechimea în specialitate etc.), psihici (educaţia, experienţa muncitorului etc.).

Factorii reprezentaţi în figura 32 constituie un ansamblu de cauze care exercită o serie de efecte asupra nivelului productivităţii muncii în industrie, influenţînd în continuare creşterea acestui indicator.

Corelaţiile statistice dintre productivitatea muncii şi factorii care o condiţionează ne arată în ce măsură indicatorii comparaţi variază împreună ca intensitate, adică creşterea unuia este însoţită sau nu de creşterea sau descreşterea celuilalt. Numai efectul exercitat de ansamblul

Scanned with OKEN Scanner

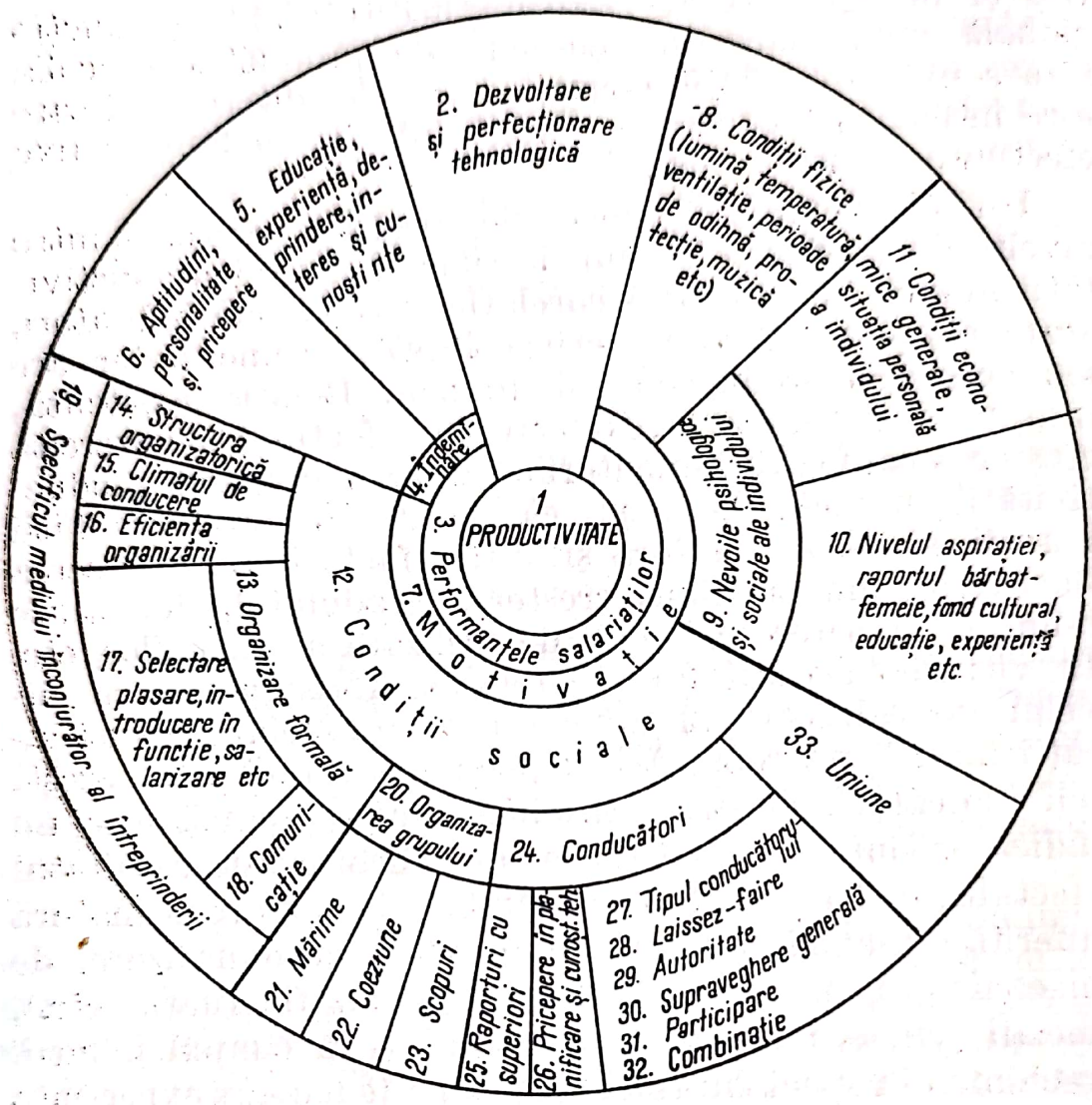


Fig. 32. — Factorii majori care afectează productivitatea muncii salariaților [5]

factorilor din figura 32 contribuie la atingerea unui anumit nivel al indicatorului productivității muncii. În acest caz, fiecare factor în parte exercită un efect parțial asupra acestui indicator sintetic.

Exemplele de mai sus evidențiază faptul că o corelație statistică poate oferi sau nu o informație oarecare asupra inferenței cauzale existente între două sau mai multe

variabile. De aceea, o simplă evidenţiere a corelaţiei statistice nu procură informaţii sigure pentru studiul inferenţelor de la cauză la efect. Datorită acestui fapt se apelează deseori la noţiunea de *relaţie cauzală*. În mod frecvent, printr-o *relaţie cauzală* se înţelege o dependenţă demonstrabilă între două variabile, din care cel puţin una este *cauza* unui *efect* exercitat asupra celeilalte. Deoarece corelaţiile statistice între variabile constituie proba relaţiilor cauzale, discuţiile asupra dependenţei acestor relaţii sînt oportune doar în prezenţa unor corelaţii statistice demonstrate. De exemplu, dacă nu va exista nici o corelaţie statistică între suprafaţa agricolă ce revine în medie pe un locuitor dintr-o unitate teritorial-administrativă oarecare (judeţ) şi ponderea muncitorilor în totalul populaţiei din zonă, va lipsi orice relaţie cauzală între aceşti doi indicatori.

Deducerea relaţiilor cauzale din corelaţiile statistice este deosebit de dificilă, datorită greutăţilor generate de eliminarea efectelor variabilelor care pot produce relaţii cauzale aparente, false. Aceste dificultăţi pot fi însă învinse prin construirea unor modele matematice de inferenţe cauzale între variabile, modele care să permită eliminarea efectelor aparente şi măsurarea relaţiilor de dependenţă reale [3].

În studiul fenomenelor şi proceselor sociale cu ajutorul metodelor şi modelelor matematice, relaţiile cauzale au o valoare mult mai mare decît a corelaţiilor statistice, deoarece primele stabilesc o dependenţă pe cînd celelalte o asociere între variabile.

Întrucît relaţiile cauzale pot fi deduse din corelaţiile statistice, construcţia modelelor matematice de inferenţe cauzale va trebui să se bazeze atît pe corelaţiile statistice cît şi pe relaţiile cauzale. În structura acestor modele matematice, corelaţiile statistice vor constitui proba relaţiilor cauzale, deoarece inexistenţa unor asocieri între variabile exclude existenţa unor inferenţe cauzale între acestea [6].

Scanned with OKEN Scanner

4.3. DEDUCEREA RELAȚIILOR CAUZALE DIN RELAȚIILE STATISTICE

Studiul relaţiilor dintre două variabile se prezintă destul de simplu atît timp cît ne mulţumim cu stabilirea intensităţii asociaţiei între aceste variabile. Dacă, însă, dorim să facem inferenţe de la cauză la efect, apar dificultăţi serioase. În general, pentru a interpreta o relaţie cauzală dintre două variabile, adică pentru a face inferenţe de la cauză la efect, trebuie să ţinem seama uneori de una sau mai multe variabile de control (C). Atunci cînd nu putem pune în evidenţă o relaţie cauzală care să nu fie influenţată de alte variabile, trebuie să recurgem la o analiză multivariată, deoarece aceasta poate identifica o structură socială în întreaga ei complexitate [7].

Pentru a analiza în profunzime structura unui proces social dat va trebui ca din ansamblul de informaţii existente să desprindem variabilele explicite şi să studiem corelaţiile statistice existente între aceste variabile. Desigur că, în acest caz, corelaţiile statistice vor participa într-o măsură importantă la realizarea inferenţelor cauzale şi invers, relaţiile cauzale ne vor sugera corelaţiile existente între variabilele respective. Deoarece există o interdependenţă între corelaţiile statistice şi relaţiile cauzale, identificarea influenţelor cauzale poate fi făcută prin evidenţierea concretă a asocierii existente între două variabile. Cu alte cuvinte, stabilirea corelaţiilor statistice poate servi la deducerea relaţiilor cauzale [8], [9].

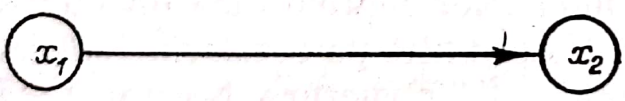


Fig. 33. — Structura unei relaţii simple între două variabile

Spre exemplificare, vom considera că variabila x_1 produce un efect asupra variabilei x_2, această relaţie putîndu-se reprezenta simbolic cu ajutorul unui arc şi a două noduri (puncte), ca în figura 33.

Scanned with OKEN Scanner

În acest caz, relaţia cauzală este asimetrică, adică efectul unei variabile x_1 asupra variabilei x_2 este dirijat într-o singură direcţie. Această relaţie cauzală poate fi reprezentată, în mod convenţional, cu ajutorul matricii 11.

Efectul \ Cauza	x_1	x_2
x_1	1	0
x_2	1	1

Matricea 11. — Matricea asociată unei relaţii cauzale între două variabile

În această matrice, o legătură de cauză-efect am indicat-o prin 1 iar inexistenţa unei astfel de legături prin 0; variabilele ce se influenţează pe ele înseşi le-am indicat tot prin 1.

Ideea de a utiliza corelaţia liniară în studiul inferenţelor cauzale aparţine lui Herbert Simon [10], [11] care a folosit o serie de modele cauzale pentru testarea anumitor ipoteze de inferenţe cauzale concrete. Pentru construirea acestor modele, reprezentate prin grafice — reţea orientate şi matrici asociate lor, s-au formulat ecuaţii liniare pentru fiecare variabilă în parte, adică pentru *variabila independentă* (x_1) şi pentru *variabila dependentă* (x_2).

Într-o relaţie cauzală între două variabile, variabila independentă (x_1) va depinde numai de ea însăşi iar variabila dependentă (x_2) se va schimba în raport cu relaţiile cauzale descrise prin ecuaţii liniare. Desigur că aceste ecuaţii vor fi afectate de o serie de fenomene întîmplătoare, neprevăzute, care vor exercita o influenţă asupra relaţiilor liniare prin intermediul unui *termen de eroare* sau termen rezidual (e). De fapt, acest termen exprimă

Scanned with OKEN Scanner

acțiunea factorilor care nu apar în mod explicit în structura studiată [9].

Utilizarea termenului rezidual (e) la formularea modelelor unor structuri de relații cauzale între două variabile (x_1 și x_2) ne sugerează faptul că previziunile făcute cu ajutorul acestor modele pot fi doar probabiliste. Dacă însă, vom reuși să eliminăm în mod teoretic, valorile termenilor de eroare, fenomenele întîmplătoare vor putea fi eliminate și prin urmare, vom putea produce exact rezultatul prezis. În acest caz, ecuațiile liniare vor căpăta semnificație de operare, de manipulare. Datorită acestui fapt, asemenea ecuații au fost considerate de Simon drept *mecanisme* liniare [11]. De exemplu, pentru structura cauzală simplă din figura 33 putem scrie pentru variabilele x_1 și x_2 următorul sistem de ecuații liniare:

$$x_1 = e_1;$$

$$b_{12}\, x_1 + x_2 = e_2\,.$$

Acestui model matematic i se poate asocia matricea 12.

$$\begin{pmatrix} 1 & 0 \\ b_{12} & 1 \end{pmatrix}$$

Matricea 12. — Matricea coeficienților asociată structurii cauzale din fig. 33

Deoarece coeficientul lui x_1 are o corespondență precisă în matricea din fig. 33 și se presupune că x_1 precede pe x_2, în matricea coeficienților de mai sus nu mai apare termenul b_{21} care este 0.

Termenii e_1 și e_2, indicați pentru fiecare ecuație în parte sînt termenii reziduali. Se presupune că sînt necorelați unul cu altul, deoarece se admite că nu există o altă variabilă (C) care să afecteze atît pe x_1 cît și pe x_2. Făcînd ipoteza că termenii reziduali (e) sînt necorelați am admis

Scanned with OKEN Scanner

faptul că produsul abaterilor normate (standard) a termenilor e_1 şi e_2 este 0, adică [8]:

$$r_{e_1 e_2} = \frac{1}{N} \sum_{i=1}^{N} \frac{e_1 \cdot e_2}{\sigma_{e_1} \cdot \sigma_{e_2}} = 0,$$

unde:

N — numărul subiecţilor anchetaţi;
σ_{e_1} — abaterea medie pătratică a erorilor e_1;
σ_{e_2} — abaterea medie pătratică a erorilor e_2.

Efectuînd o serie de simplificări, obţinem:

$$\sum_{i=1}^{N} e_1 \cdot e_2 = 0$$

Formulînd cea de a doua ipoteză, S i m o n a fundamentat metoda sa de multiplicare a două ecuaţii [8]:

$$b_{12}\, x_1^2 + x_1\, x_2 = e_1\, e_2 .$$

Însumînd această ecuaţie şi făcînd ca produsul termenilor reziduali să fie 0, putem scrie:

$$b_{12} \Sigma\, x_1^2 + \Sigma\, x_1\, x_2 = e_1\, e_2 = 0 .$$

De aici rezultă valoarea coeficientului b_{12} care reflectă relaţia cauzală dintre x_1 şi x_2, precum urmează:

$$b_{12} = -\frac{\Sigma\, x_1\, x_2}{\Sigma\, x_1^2} = -\, r_{12} \frac{\sigma_1}{\sigma_2},$$

unde:

r_{12} — coeficientul de corelaţie statistică dintre variabilele x_1 şi x_2;
σ_1 — abaterea medie pătratică a valorilor variabilei x_1;
σ_2 — abaterea medie pătratică a valorilor variabilei x_2.

Scanned with OKEN Scanner

Rezultă că acest coeficient b_{12} va fi 0 sau nonzero după cum r_{12} va fi 0 sau nonzero. Deci, coeficientul de corelaţie r_{12} este proba relaţiei cauzale reflectate de coeficientul b_{12}.

Descriere	Diagramă	Ecuaţii liniare	Previziune
Relaţie cauzală (de cauză-efect)		$x_1 = e_1$ $b_{21}\, x_1 +$ $+ x_2 = e_2$	$b_{12}\, x_1^2 + x_1\, x_2 = e_1\, e_2$ $b_{12} \sum x_1^2 + \sum x_1\, x_2 = \sum e_1\, e_2 = 0$ $b_{12} = -\dfrac{\sum x_1\, x_2}{\sum x_1^2} = -r_{12} \cdot \dfrac{\sigma_1}{\sigma_2}$

Fig. 34. — Modelul matematic al relaţiei cauzale existente între o variabilă independentă (x_1) şi una dependentă (x_2)

Examinînd modelul matematic al structurii cauzale din fig. 34, putem trage concluzia că relaţiile cauzale pot fi deduse din corelaţiile statistice [4]. Desigur că, în cazul unei singure legături de cauză-efect, interpretarea cauzală a unei corelaţii statistice este relativ uşoară. Dacă numărul de variabile şi de relaţii cauzale creşte acestea din urmă pot fi deduse din corelaţiile statistice, însă cu mare dificultate [1].

În cazul unui model cauzal cu trei variabile (x_1, x_2, x_3) care se interacţionează, graful orientat poate fi reprezentat ca în fig. 35.

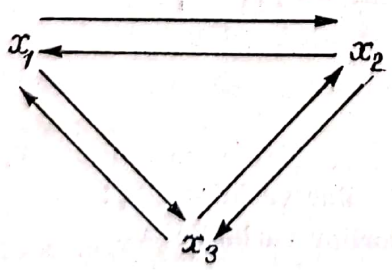


Fig. 35. — Structura unei relaţii cauzale între trei variabile

Scanned with OKEN Scanner

Interacţiunea între aceste trei variabile poate fi pusă în evidenţă cu ajutorul matricii 13 [7].

Efectul \ Cauza	x_1	x_2	x_3
x_1	1	1	1
x_2	1	1	1
x_3	1	1	1

Matricea 13. — Matricea asociată unei structuri de relaţii cauzale între trei variabile

Pentru structura cauzală din figura 35, sistemul de ecuaţii liniare se poate scrie astfel [12]:

$$x_1 + b_{21}x_2 + b_{31}x_3 = e_1;$$
$$b_{12}x_1 + x_2 + b_{32}x_3 = e_2;$$
$$b_{13}x_1 + b_{23}x_2 + x_3 = e_3.$$

În acest sistem de ecuaţii apar ca necunoscute cei şase coeficienţi b_{ij} şi cei trei termeni reziduali e_j. Sistemul de ecuaţii liniare poate fi rezolvat dacă emitem ipoteza că termenii reziduali sînt necorelaţi şi dacă formulăm 5 proprietăţi discriminatorii ale dublei săgeţi. Presupunînd că fiecare din aceste 5 modele cauzale derivate este constituit din trei variabile, astfel încît între perechile de variabile să existe cel mult o singură legătură cauzală, putem alcătui modelul matematic al întregii structuri cauzale, ca în figura 36 .

Pentru fiecare model derivat putem scrie ecuaţiile liniare ale sistemului cauzal general din figura 36 după cum urmează [1]:

a) *Cauza dublă.* Făcînd ipoteza că erorile reziduale nu sînt corelate, suma produselor erorilor reziduale se poate considera ca fiind egală cu 0, adică :

$$\Sigma e_1 e_2 = \Sigma e_1 e_3 = \Sigma e_2 e_3 = 0.$$

Scanned with OKEN Scanner

În acest caz, schemei cauzale duble derivate din figura 36 i se poate asocia următorul sistem de ecuații liniare:

$$b_{12}\sum x_1^2 + \sum x_1x_2 + b_{32}\sum x_1x_3 = \sum e_1e_2 = 0,$$
$$b_{12}\sum x_1x_3 + \sum x_2x_3 + b_{32}\sum x_3^2 = \sum e_2e_3 = 0,$$
$$\sum x_1x_3 = \sum e_1e_3 = 0,$$

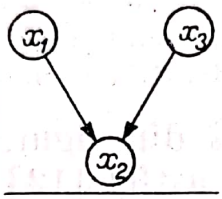
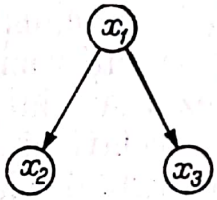
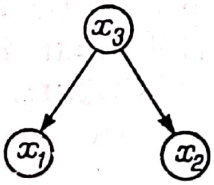
Nr. crt.	Descriere	Diagramă	Ecuații liniare	Ipoteză
1	Cauză dublă		$x_1 = e_1$ $b_{12}x_1 + x_2 + b_{32}x_3 = e_2$ $x_3 = e_3$	Dacă $r_{23} \neq 0$ și $r_{21} \neq 0$ obținem: $r_{31} = 0$
2	Secvențe de dezvoltare		$x_1 + b_{31}x_3 = e_1$ $b_{12}x_1 + x_2 = e_2$ $x_3 = e_3$	Dacă $r_{13} \neq 0$ și $r_{21} \neq 0$ obținem: $r_{23} = r_{13} \cdot r_{21}$
3	Efect dublu		$x_1 = e_1$ $b_{12}x_1 + x_2 = e_2$ $b_{13}x_1 + x_3 = e_3$	Dacă $r_{21} \neq 0$ și $r_{31} \neq 0$, obținem: $r_{32} = r_{21} \cdot r_{31}$
4	Corelație falsă		$x_1 + b_{31}x_3 = e_1$ $x_2 + b_{32}x_3 = e_2$ $x_3 = e_3$	Dacă $r_{13} \neq 0$ și $r_{23} \neq 0$, obținem: $r_{21} = r_{13} \cdot r_{23}$
5	Secvență de intervenție		$x_1 + b_{13}x_3 = e_1$ $x_2 + b_{32}x_3 = e_2$ $x_3 = e_3$	Dacă $r_{31} \neq 0$ și $r_{23} \neq 0$ obținem: $r_{21} = r_{31} \cdot r_{23}$

Fig. 36. — Modelul matematic al relațiilor cauzale dintre trei variabile interdependente (x_1, x_2, x_3)

Împărţind prima şi ultima ecuaţie la $\frac{\Sigma x_1^2}{N}$ iar cea de a doua la $\frac{\Sigma x_3^2}{N}$, obţinem o serie de relaţii între coeficienţii b_{ij} şi coeficienţii de corelaţie r_{ij} :

$$b_{12} + r_{12} + b_{32}r_{13} = 0;$$

$$b_{12}r_{13} + r_{23} + b_{32} = 0,$$

$$r_{13} = 0.$$

b) *Secvenţe de dezvoltare.* Utilizînd schema cauzală derivată a secvenţei de dezvoltare din figura 36, putem scrie sistemul de ecuaţii liniare :

$$\Sigma x_1x_3 + b_{31}\Sigma x_3^2 = \Sigma e_1e_3 = 0,$$

$$b_{12}\Sigma x_1x_3 + \Sigma x_2x_3 = \Sigma e_2e_3 = 0,$$

$$b_{12}\Sigma x_1^2 + \Sigma x_1x_2 + b_{31}b_{12}\Sigma x_1x_3 + b_{31}\Sigma x_2x_3 = \Sigma e_1e_2 = 0.$$

Împărţind prima ecuaţie la $\frac{\Sigma x_3^2}{N}$, cea de a doua şi a treia la $\frac{\Sigma x_1^2}{N}$, obţinem :

$$r_{13} + b_{31} = 0,$$

$$b_{12}r_{31} + r_{32} = 0,$$

$$b_{12} + r_{12} + b_{31}b_{12}r_{31} + b_{31}r_{23} = 0,$$

Scanned with OKEN Scanner

c) *Efect dublu.* Avînd în vedere schema cauzală derivată a efectului dublu din figura 36, putem formula sistemul de ecuaţii liniare :

$$b_{13}\Sigma x_1^2 + \Sigma x_1 x_3 = \Sigma e_1 e_3 = 0,$$

$$b_{12}b_{13}\,\Sigma x_1^2 + b_{13}\,\Sigma x_1 x_2 + b_{12}\,\Sigma x_1 x_3 + \Sigma x_2 x_3 = \Sigma\, e_2 e_3 = 0,$$

$$b_{12}\,\Sigma x_1^2 + \Sigma x_1 x_2 = \Sigma\, e_1 e_2 = 0$$

Împărţind fiecare ecuaţie la $\frac{\Sigma x_1^2}{N}$, obţinem :

$$b_{13} + r_{13} = 0,$$

$$b_{12}b_{13} + b_{13}r_{12} + b_{12}r_{13} + r_{23} = 0,$$

$$b_{12} + r_{12} = 0.$$

d) *Corelaţia falsă.* Folosind schema cauzală derivată a corelaţiei false putem scrie următorul sistem de ecuaţii liniare :

$$\Sigma x_1 x_3 + b_{31}\,\Sigma x_3^2 = \Sigma e_1 e_3 = 0,$$

$$\Sigma x_2 x_3 + b_{32}\,\Sigma x_3^2 = \Sigma e_2 e_3 = 0.$$

$$\Sigma x_1 x_2 + b_{32}\,\Sigma x_1 x_3 + b_{31}\Sigma\, x_2 x_3 + b_{31}\,b_{32}\,\Sigma\, x_{32} = \Sigma e_1 e_2 = 0.$$

Împărţind fiecare ecuaţie la $\frac{\Sigma x_3^2}{N}$, obţinem :

$$r_{13} + b_{31} = 0,$$

$$r_{23} + b_{32} = 0,$$

$$r_{12} + b_{32}r_{13} + b_{31}r_{23} + b_{31}b_{32} = 0,$$

Scanned with OKEN Scanner

e) Secvență de intervenție. Utilizînd schema cauzală derivată a secvenței de intervenție din figura 36, putem scrie sistemul de ecuații liniare:

$$\Sigma x_1 x_3 + b_{31} \Sigma x_3^2 = \Sigma e_1 e_2 = 0,$$

$$\Sigma x_2 x_3 + b_{32} \Sigma x_3^2 = \Sigma e_2 e_3 = 0,$$

$$\Sigma x_1 x_2 + b_{31} \Sigma x_2 x_3 + b_{32} \Sigma x_1 x_3 + b_{31} b_{32} \Sigma x_3^2 = \Sigma e_1 e_2 = 0.$$

Împărțind fiecare ecuație la $\frac{\Sigma x_3^2}{N}$, obținem:

$$r_{13} + b_{31} = 0,$$

$$r_{23} + b_{32} = 0,$$

$$r_{12} + b_{31} r_{23} + b_{32} r_{13} + b_{31} b_{32} = 0.$$

În concluzie, sistemelor cauzale derivate de mai sus li se pot asocia sisteme de relații matematice dintre coeficienții de corelație (r_{ij}) și coeficienții care pun în evidență inferențele cauzale (b_{ij}).

Procedînd în același mod, se pot construi modele cauzale multivariate (care includ multe variabile) [19].

4.4. ANALIZA DEPENDENȚEI ÎNTRE DOUĂ VARIABILE

Influențele directe ale unei variabile independente (x) asupra alteia dependente (y) formează serii de repartiție bidimensionale, care de obicei sînt corelate simplu. După forma legăturii dintre variabile, corelația simplă poate fi curbilinie sau liniară, în funcție de dispersia datelor

Scanned with OKEN Scanner

obținute din observații. Corelația curbilinie se exprimă sintetic cu ajutorul ecuației unei curbe (parabolă, hiperbolă, curbă exponențială etc.) iar cea liniară prin ecuația unei drepte.

În analiza dependenței dintre două variabile (x și y) se utilizează frecvent noțiunea de *linie de regresie* [14]. Linia de regresie teoretică a unei serii de repartiție bidimensionale reprezintă linia care urmărește mijloacele intervalelor medii pe perechile de puncte din diagrama de dispersie. Această linie poate avea forme diferite : o dreaptă sau o curbă oarecare.

În practică se pot întîlni două categorii de linii de regresie între două variabile corelate :

a) linia de regresie a lui y spre x, adică linia obținută în ipoteza că valorile lui x sînt exacte ;

b) linia de regresie a lui x spre y, adică linia obținută în ipoteza că valorile lui y sînt exacte.

În graficul din figura 37 se dă un exemplu de reprezentare a liniei de regresie a lui y spre x [15].

Graficul din figura 38 constituie un exemplu de reprezentare a liniei de regresie a lui x spre y.

Orice funcție continuă, oricît de complicată ar fi, poate fi reprodusă cu precizie, într-un interval dat, apriori, printr-un polinom de gradul n, adică prin expresia [4] :

$$y = a_{12}x + b_{12}x^2 + c_{12}x^3 + \ldots + z_{12}x^m + e_2.$$

Cu ocazia emiterii ipotezei asupra formei curbei de regresie putem utiliza expresia acestui polinom. În practică însă, dificultățile implicate de alegerea formei curbei de regresie pot fi eliminate prin introducerea unei ipoteze simplificatoare, constînd din reprezentarea geometrică a unei funcții printr-o dreaptă în plan, adică prin ecuația[1] :

$$y = a_{12}\,x + e_2\,.$$

[1] În acest caz, variabila $x = e_1$.

Pentru un „nor de puncte" care reprezintă distribuţia datelor statistice pentru y în raport cu x, se pot trasa mai multe drepte. Dintre toate aceste drepte posibile, trebuie

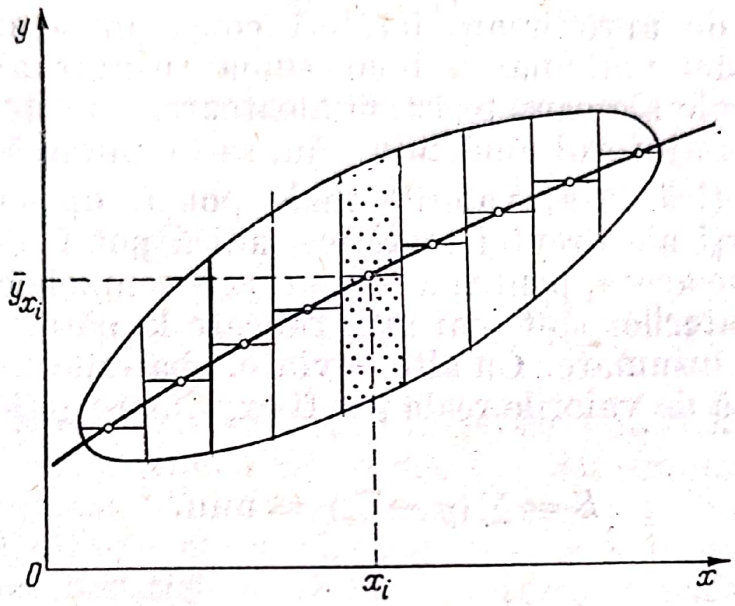


Fig. 37. — Linia de regresie a lui y spre x

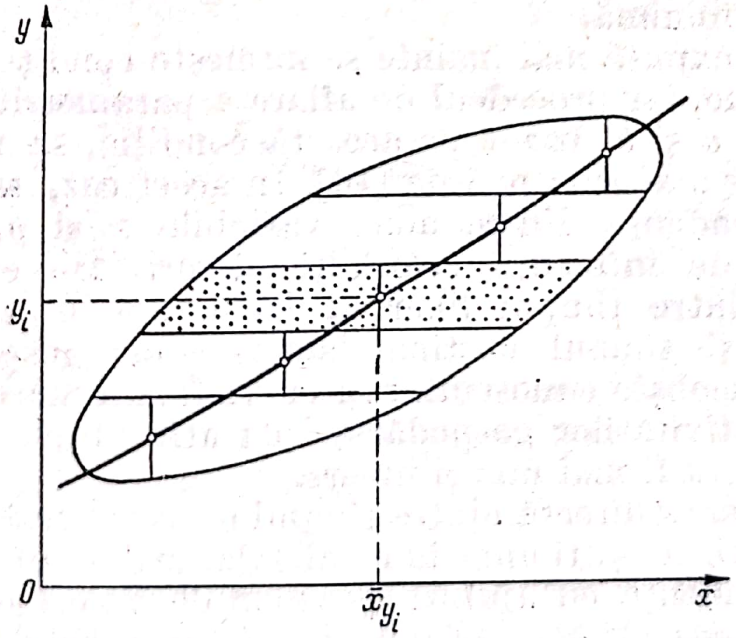


Fig. 38. — Linia de regresie a lui x spre y

Scanned with OKEN Scanner

să alegem una care să corespundă, cel mai bine, datelor noastre statistice. În acest scop, trebuie să determinăm coeficienţii a_{12} şi e_2 din ecuaţia de mai sus, astfel încît aceasta să definească dreapta unică care se apropie, mai mult ca oricare alta, de punctele considerate. Stabilirea criteriului de aproximare implică construirea unui asemenea model matematic, încît suma tuturor abaterilor valorilor reale ale caracteristicii aleatoare, faţă de valorile calculate cu ajutorul modelului său, să fie minimă.

În practică însă, valorile reale pot fi dispersate de ambele părţi ale dreptei teoretice, adică pot fi de semne contrare. De aceea, pentru a fi înlăturate semnele contrare, valorile abaterilor sînt mai întîi ridicate la pătrat şi după aceea sînt însumate. Cu alte cuvinte, abaterile modelului teoretic faţă de valorile reale pot fi exprimate astfel :

$$S = \Sigma (y - \overline{y}_x)^2 = \min.$$

Această expresie arată că suma pătratelor abaterilor reale ale caracteristicii pe care o studiem (y), faţă de valorile calculate ($\overline{y}_x$) cu ajutorul modelului său teoretic trebuie să fie minimă.

Regula expusă mai înainte se numeşte condiţia pătratelor minime, iar procedeul de aflare a parametrilor liniei de regresie a şi b, bazat pe această condiţie, se numeşte *metoda celor mai mici pătrate* [16]. În acest caz, se analizează dependenţa dintre două variabile x şi y, făcînd abstracţie de influenţa celorlalţi factori. De exemplu, legătura dintre timpul liber al bărbaţilor în zilele de lucru (y) şi timpul destinat activităţilor gospodăreşti (x) este îndeobşte cunoscută : cu cît va fi mai mare timpul destinat activităţilor gospodăreşti cu atît timpul liber al salariaţilor va fi mai mic şi invers.

Dependenţa directă dintre timpul destinat activităţilor gospodăreşti (x) şi timpul liber al salariaţilor (y) poate fi pusă în evidenţă cu ajutorul seriilor de valori ordonate după mărimea primei variabile (x), ca în tabelul 16.

Scanned with OKEN Scanner

Tabelul 16

Categoria de salariaţi	Timpul destinat activităţilor gospodăreşti (x)	Timpul liber (y)	x^2	y^2	x, y
0	1	2	3	4	5
Ingineri şi tehnicieni	0,39	3,32	0,15	11,02	1,29
Funcţionari	0,46	3,18	0,21	10,16	1,46
Muncitori	0,59	3,02	0,35	9,12	1,78
Personal de serviciu	1,46	2,15	2,13	4,62	3,14
T o t a l	2,90	11,67	2,84	39,92	7,67

Într-adevăr, examinînd diferitele valori ale timpului liber în cursul unei singure zile (y) corespunzătoare timpului destinat activităţilor gospodăreşti (x), se remarcă tendinţa de creştere a caracteristicii x o dată cu descreşterea caracteristicii y. Dacă reprezentăm grafic perechile de valori concrete într-un sistem de coordonate rectangulare, legătura dintre cele două variabile se poate reprezenta printr-o linie frîntă cu tendinţă de scădere, ca în figura 39.

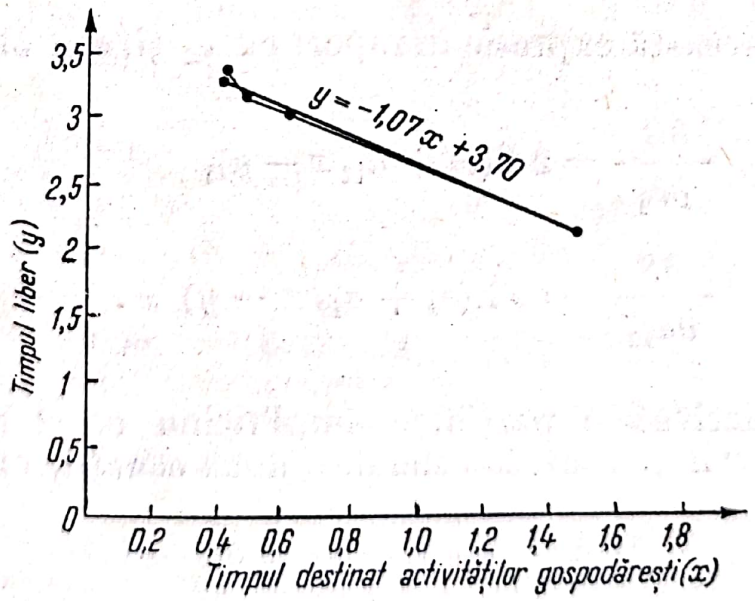


Fig. 39. — Reprezentarea grafică a liniei de regresie

Inflexiunile liniei de regresie din figura 39 reflectă influenţa altor factori asupra timpului liber, în afară de timpul destinat activităţilor gospodăreşti. Ca să eliminăm influenţa acestor factori se trece la ajustarea valorilor y în ipoteza că între y şi x există o dependenţă funcţională. În cazul unei corelaţii liniare dintre aceste variabile, la ajustare, valorile efective y se înlocuiesc cu valorile calculate cu ajutorul relaţiei $y = a_{12}x + e_2$ unde a_{12} şi e_2 reprezintă parametrii liniei de regresie liniară. În această relaţie, valoarea variabilei dependente sau rezultative (y) se exprimă în funcţie de valoarea variabilei independente sau a argumentului (x), toate celelalte variabile fiind considerate drept mărimi constante şi fiind exprimate prin parametrii a_{12} şi e_2. În acelaşi timp, toţi factori care determină pe y, în afară de x, sînt mărimi medii. De aceea, valorile ajustate ale lui y sînt de asemenea mărimi medii.

Pentru a fi posibilă determinarea valorilor y trebuie să determinăm în prealabil parametrii liniei de regresie (a_{12} şi e_2). Aceşti parametri pot fi calculaţi cu ajutorul metodei celor mai mici pătrate, care satisface condiţia [16]:

$$S = \Sigma (e_2 + a_{12}\, x - y)^2 = \text{minim}\,.$$

Derivînd această expresie în raport cu e_2 şi a_{12}, obţinem:

$$\frac{\partial S}{\partial e_2} = 2\,\Sigma(e_2 + a_{12}\, x - y),$$

$$\frac{\partial S}{\partial a_{12}} = 2\,\Sigma(e_2 + a_{12}\, x - y)\; x\,.$$

Anulînd derivatele parţiale, simplificînd cu 2 şi introducînd pe x în paranteza celei de a doua ecuaţii, obţinem:

$$\Sigma\,(e_2 + a_{12}\, x - y) = 0\,,$$

$$\Sigma\,(e_2\, x + a_{12}x^2 - xy) = 0\,.$$

Scanned with OKEN Scanner

Sistemul de ecuaţii normale poate fi obţinut prin descompunerea sumelor în termeni şi prin scoaterea parametrilor a_{12} şi e_2 în afara semnului sumei :

$$ne_2 + a_{12}\,\Sigma x = \Sigma y,$$

$$e_2\,\Sigma\,x + a_{12}\,\Sigma\,x^2 = \Sigma\,x\,y.$$

Înlocuind acum datele iniţiale şi cele calculate în tabelul 1, obţinem :

$$4e_2 + 2{,}90\;a_{12} = 11{,}67,$$

$$2{,}90e_2 + 2{,}84\;a_{12} = 7{,}67.$$

Rezolvînd acest sistem cu ajutorul determinanţilor vom obţine :

$$e_2 = \frac{\begin{vmatrix} 11{,}67 & 2{,}90 \\ 7{,}67 & 2{,}84 \end{vmatrix}}{\begin{vmatrix} 4 & 2{,}90 \\ 2{,}90 & 2{,}84 \end{vmatrix}} = \frac{11{,}67 \times 2{,}84 - 7{,}67 \times 2{,}90}{4 \times 2{,}84 - 2{,}90 \times 2{,}90} =$$

$$= \frac{10{,}90}{2{,}95} = 3{,}70$$

$$a_{12} = \frac{\begin{vmatrix} 4 & 11{,}67 \\ 2{,}90 & 7{,}67 \end{vmatrix}}{\begin{vmatrix} 4 & 2{,}90 \\ 2{,}90 & 2{,}84 \end{vmatrix}} = \frac{4 \times 7{,}67 - 11{,}67 \times 2{,}90}{4 \times 2{,}84 - 2{,}90 \times 2{,}90} =$$

$$= \frac{-3{,}16}{2{,}95} = -1{,}07$$

Scanned with OKEN Scanner

Inlocuind valorile parametrilor a_{12} şi e_2 în relaţia dintre bugetul de timp liber (y) şi timpul destinat activităţilor gospodăreşti (y), obţinem :

$$y = a_{12}x + e_2 = -1{,}07\,x + 3{,}70.$$

Parametrii a_{12} şi e_2 ai ecuaţiei de regresie permit calcularea mărimii medii a lui y pentru un x dat, celelalte condiţii fiind considerate identice.

În sens geometric, parametrul a_{12}, care se numeşte coeficient de regresie, reprezintă panta liniei drepte, adică : $a_{12} = \text{tg}\alpha$, unde α este unghiul făcut de dreaptă cu axa x.

Coeficientul de regresie a_{12} stabileşte gradul de proporţionalitate dintre mărimea variabilei cercetate şi rezultatul obţinut. El arată măsura în care se schimbă variabila y cînd x se schimbă cu o unitate. În cazul nostru, dacă timpul destinat activităţilor gospodăreşti (x) creşte cu 1 oră, timpul liber al salariaţilor va scădea cu 1,07 ore, deoarece coeficientul a_{12} are semnul negativ în ecuaţia de regresie liniară $y = -1{,}07\,x + 3{,}70$.

Coeficientul de regresie a_{12} capătă valoare pozitivă în cazul corelaţiei directe şi valoare negativă în cazul corelaţiei inverse. Cu alte cuvinte, linia de regresie în cazul corelaţiei directe exprimă o tendinţă de creştere (întrucît a_{12} este pozitiv), iar în cazul corelaţiei inverse, o tendinţă de scădere (întrucît a_{12} are sens negativ).

Utilizînd metoda celor mai mici pătrate, coeficientul de regresie a lui x spre y îl putem determina şi cu ajutorul formulei [1] :

$$a_{12} = r_{12}\,\frac{\sigma_2}{\sigma_1},$$

unde :

r_{12} — coeficientul de corelaţie liniară între x şi y;
σ_1;σ_2 — abaterile medii pătratice ale variabilelor x şi y.

Scanned with OKEN Scanner

Coeficientul de regresie a lui y spre x va fi egal cu $r_{12}\frac{\sigma_1}{\sigma_2}$. Dacă coeficientul de corelaţie r_{12} va fi nul, atunci coeficientul de regresie a_{12} va fi de asemenea nul.

Coeficientul e_2 poate fi interpretat ca o eroare reziduală care nu depinde de variaţia variabilei analizate : uneori se consideră că reflectă influenţa pe care o exercită asupra procesului studiat, variabilele de care am făcut abstracţie şi a căror acţiune nu depinde de variabila cercetată ; alteori nu i se dă nicio semnificaţie independentă, ci una de calcul[2] [4].

În virtutea celor arătate mai sus, determinarea pantei liniei de regresie liniară constituie una din problemele principale ale dependenţei dintre două variabile. De multe ori, cunoaşterea în profunzime a unei astfel de dependenţe poate constitui un element esenţial pentru alegerea corectă a formei liniei de regresie. Această alegere a liniei de regresie poate fi soluţionată în condiţii optime numai pentru o dependenţă între doi indicatori din care nici unul nu este sintetic.

Atunci cînd se stabileşte o dependenţă dintre un indicator analitic şi unul sintetic, putem obţine rezultate care pot fi influenţate de erori inacceptabile pentru cercetarea întreprinsă. De pildă, chiar bugetul de timp al salariaţilor poate fi privit ca un indicator sintetic, deoarece asupra lui poate acţiona un ansamblu de factori : timpul destinat activităţilor gospodăreşti, timpul de deplasare la locul de muncă, timpul de efectuare a cumpărăturilor etc.

Aportul cercetătorului poate fi deosebit de util într-o activitate de analiză matematică concretă, atunci cînd acesta poate deosebi variabilele sintetice de cele analitice, astfel încît să sesizeze aportul variabilelor care generează o serie de relaţii de dependenţă aparent false.

---

[2] În cazul seriilor dinamice, coeficientul e_2 poate reprezenta nivelul iniţial al seriei.

Efectul variabilelor „perturbatorii" în studiul inferenţelor cauzale simple (între două variabile) trebuie eliminat din analiză. Desigur că eliminarea efectului variabilelor care generează dependenţe aparente este extrem de dificilă, deoarece în practică există relaţii numeroase între variabile; cercetarea doar a unor relaţii cauzale izolate din mulţimea relaţiilor existente în structura unui proces social oarecare este destul de riscantă dacă avem în vedere faptul că relaţiile acestea pot fi influenţate cauzal de celelalte variabile din sistem [9].

Deoarece în domeniul sociologiei, structurile cauzale simple sînt foarte rare, analiza relaţiilor de dependenţă între două variabile poate fi efectuată numai atunci cînd facem concesii mari în privinţa erorilor ce apar în studiul inferenţelor cauzale.

Prin urmare, analiza dependenţei dintre două variabile trebuie făcută cu multă atenţie, avînd în vedere riscurile de „evidenţiere" a unor inferenţe cauzale aparente, de suprafaţă. De aceea, este de preferat să analizăm inferenţele cauzale multiple, proprii de altfel structurilor cauzale complexe prin excelenţă.

O analiză matematică multivariată a structurilor cauzale este superioară în mod evident, deoarece exclude în mare măsură, posibilitatea ca dependenţele dintre variabile să fie „alterate" de alte influenţe cauzale exterioare.

4.5. COEFICIENTUL DE DEPENDENŢĂ A LUI WRIGHT

Un model cauzal de tipul „x este cauza lui y" se poate traduce într-un sistem de ecuaţii liniare [1]:

$$x = e_1,$$

$$y = a_{12}\, x + e_2 .$$

În acest caz, din mulțimea de variabile între care există relații de dependență s-au ales doar două variabile (x și y), făcînd abstracție de influența celorlalte variabile.

Desigur că această relație implică existența unei singure cauze și a unui singur efect. De aceea, putem spune că y depinde de x, dar x nu depinde de y. Cu alte cuvinte, y este o funcție de x, ceea ce se poate scrie astfel [4]:

$$y = f(x)$$

În anumite condiții, modelul cauzal de tipul „x este cauza lui y" este insuficient definit. De pildă, dacă variabila x este efectul unui ansamblu cauzal x, compus din valori particulare x_1, x_2 . . ., x_n, atunci putem spune că: „x este doar una din cauzele lui y" [4].

Prin urmare, în condițiile pluralității cauzelor lui y, cazul cauzei unice nu poate fi decît un caz limită. În condițiile în care „x este doar una din cauzele lui y" va trebui să reformulăm sistemul de ecuații liniare de mai înainte. În acest caz, va trebui să ținem seama de întregul ansamblu cauzal al lui y [2].

Analiza dependenței implică emiterea ipotezei că asupra variabilei y exercită efect doar o singură variabilă x și nu un ansamblu de variabile x_j, adică: x_1, x_2, . . . x_n. Numai pentru o structură simplă și elementară în care „o variabilă y depinde liniar de o variabilă x" putem scrie relația: $y = a_{12}\, x + e_2$, unde coeficientul de regresie este $a_{12} = r_{12} \dfrac{\sigma_2}{\sigma_1}$. Coeficientul de corelație r_{12} măsoară proporția variației abaterii medii pătratice (σ_i) a celor două variabile; cantitatea $a_{12}^2 \dfrac{\sigma_1^2}{\sigma_2^2}$ poate fi interpretată ca o măsură a influenței cauzale a unei variabile asupra alteia[3].

[3] O primă interpretare a coeficientului de dependență a fost dată de Wright (1934). De aceea, uneori se utilizează expresia „coeficientul lui Wright".

De aceea, expresia : $d_{12} = a_{12}\frac{\sigma_1^2}{\sigma_2^2}$ poartă denumirea de coeficient de dependenţă [8]. În general, coeficientul de dependenţă poate fi interpretat ca un raport între coeficienţii de determinare (r_i) ai celor două variabile (y şi x). În baza acestei interpretări putem scrie [17] :

$$d_{12}^2 = \frac{r_2^2}{r_1^2} = \frac{\frac{\sigma_{2i}^2}{\sigma_2^2}}{\frac{\sigma_{1i}^2}{\sigma_1^2}},$$

unde : $\sigma_{2i}^2 = \frac{1}{N}\sum_{i=1}^{N}(\hat{Y}_i - \overline{Y}_i)^2$;

$$\sigma_2^2 = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \overline{Y}_i)^2;$$

$$\sigma_{1i}^2 = \frac{1}{N}\sum_{i=1}^{N}(\hat{X}_i - \overline{X}_i)^2;$$

$$\sigma_1^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \overline{X}_i)^2.$$

În aceste expresii am notat cu :

r_1^2; r_2^2 — coeficienţii de determinare ai variabilelor x şi y;
σ_1^2; σ_2^2 — dispersiile totale ale variabilelor x şi y;
σ_{1i}^2; σ_{2i}^2 — dispersiile explicative ale variabilelor x şi y;
X_i; Y_i — valorile observabile ale variabilelor x şi y;
$\hat{X}_i$; $\hat{Y}_i$ — valorile explicative ale variabilelor x şi y;
$\overline{X}_i$; $\overline{Y}_i$ — valorile medii ale variabilelor x şi y;
N — numărul datelor observabile.

Scanned with OKEN Scanner

Întrucît pătratul coeficientului de regresie liniară este :

$$a_{12}^2 = \frac{\sigma_{2i}^2}{\sigma_{1i}^2},$$

expresia lui d_{12} poate fi scrisă astfel [8] :

$$d_{12} = a_{12} \frac{\sigma_1}{\sigma_2}.$$

Acest coeficient măsoară destul de corect influența cauzală a variabilei x asupra variabilei y. El exprimă de asemenea proporția dispersiei lui y explicate prin x. Coeficientul de regresie liniară a lui y spre x fiind $a_{12} = r_{12} \frac{\sigma_2}{\sigma_1}$, expresia coeficientului de dependență d_{12} devine :

$$d_{12} = r_{12}.$$

Prin urmare, coeficientul de dependență (d_{12}) este egal cu coeficientul de corelație (r_{12}) în cazul unei relații matematice dintre variabilele x și y. Cu alte cuvinte, coeficientul de corelație (r_{12}) constituie o măsură foarte corectă a dependenței cauzale dintre x și y. De pildă, coeficientul de corelație $r_{12} = 0{,}78$, dintre gradul de urbanizare al unei unități teritorial-administrative (județ) și cheltuielile pentru acțiuni social-culturale, ar măsura foarte bine dependența cauzală între acești indicatori, dacă nici unul dintre aceștia n-ar fi afectat de un ansamblu întreg de alți indicatori, cum ar fi : producția industrială, ponderea salariaților, ponderea muncitorilor, vînzările de mărfuri etc.

Întrucît, de cele mai multe ori, fiecare din variabilele x și y între care se stabilește o dependență, implică o pluralitate de cauze, măsurarea influențelor cauzale cu aju-

Scanned with OKEN Scanner

torul coeficienţilor de corelaţie este riscantă. De aceea, în asemenea situaţii trebuie să recurgem la serviciile analizei influenţelor cauzale (dependenţelor) existente în structura modelelor cauzale complexe. De altfel, elaborarea unor modele cauzale complexe este extrem de utilă analizei matematice a proceselor sociale, deoarece structurile cauzale simple se întîlnesc rar în domeniul ştiinţelor sociale.

4.6. MODELE LINIARE DE ANALIZĂ A STRUCTURILOR CAUZALE COMPLEXE

Într-o structură cauzală complexă trebuie să pornim de la principiul pluralităţii posibile a cauzelor care generează un ansamblu de efecte asupra variabilei analizate. De aceea, prin „configuraţia cauzală particulară a lui y", vom înţelege ansamblul X_j, numit „ansamblul cauzal al lui y". În acest caz, ansamblul cauzal X_j, compus din valorile particulare ale variabilelor independente $x_1, x_2, \ldots, x_n$, corespunde unei singure valori a variabilei dependente y [4].

O structură cauzală complexă poate fi formulată algebric cu ajutorul unei funcţii de forma [1]:

$$y = f(x_1, x_2, \ldots, x_n).$$

Această funcţie exprimă asimetria relaţiilor cauzale, dînd un caracter deosebit corespondenţei între variabila y şi ansamblul ei cauzal $X_i \equiv \{x_1, x_2, \ldots, x_n\}$. Caracterul deosebit al acestei corespondenţe derivă din traducerea matematică concretă a expresiei „y depinde în mod univoc de x_j, dar x_j nu depinde în mod univoc de y". Este evident că acestei expresii îi corespunde expresia asimetrică „y este funcţie de x_j". De aici rezultă de altfel caracterul echivoc al expresiei: „y depinde de x_j, dar x_j nu depinde de y".

Scanned with OKEN Scanner

Exprimarea matematică a unei structuri cauzale complexe depinde de procedeul empiric utilizat : *experimentare* sau *observaţie* [12].

O funcţie liniară de forma : $y = f(x_1, x_2, \ldots, x_n)$, în care există o independenţă perfectă între cauze, poate fi scrisă astfel [4] :

$$y = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n,$$

$$E(x_1 x_2) = E(x_1 x_3) = E(x_2 x_3) = \ldots = E(x_1 x_n) =$$

$$= E(x_2 x_n) = E(x_3 x_n) = 0,$$

unde :

$E(x_i x_j)$ — speranţa matematică.

Fiecare coeficient a_j este coeficientul de regresie liniară a lui y în raport cu fiecare x_j luat succesiv. În acest caz, un calcul de regresie multiplă revine la o succesiune de regresii simple, iar modelul matematic de mai sus exprimă o „structură simplă".

Efectuarea unui calcul de regresie multiplă se complică atunci cînd cauzele sînt parţial dependente. Desigur că existenţa relaţiilor cauzale între cauze ne conduce în domeniul „structurilor complexe" [18].

Ecuaţiile structurale definesc relaţiile funcţionale existente între variabile. De pildă, dacă considerăm sistemul compus din [4] :

$$b_{12} x_1 + x_2 = e_2,$$

$$b_{13} x_1 + b_{23} x_2 + x_3 = e_3,$$

sau în transcrierea matricială :

$$\begin{pmatrix} b_{12} & 1 & 0 \\ b_{13} & b_{23} & 1 \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} e_2 \\ e_3 \end{pmatrix},$$

Scanned with OKEN Scanner

putem înlocui valoarea variabilelor implicate lui e_2 prin $\eta_2 = e_2 + b_{32}x_3$:

$$\begin{pmatrix} b_{12} & 1 & b_{32} \\ b_{13} & b_{23} & 1 \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \eta_2 \\ e_3 \end{pmatrix}.$$

Semnificaţia unui ansamblu de ecuaţii :

$$b_{12}\, x_1 + x_2 = e_2,$$

$$b_{13}\, x_1 + b_{23}x_2 + x_3 = e_3, \text{ cu :}$$

$$E\,(e_2x_1) = 0,\ E(e_3x_1) = 0,\quad E\,(e_3e_2) = 0,$$

apare ca :

$$x_3 = a_{13}x_1 + a_{23}x_2 + e_3.$$

Această ultimă ecuaţie exprimă dependenţa liniară a lui x_3 de cele două variabile explicite x_1 şi x_2 (x_3 nefiind în întregime determinat de cele două variabile explicite) şi de o variabilă implicită e_3. Dacă se admite ipoteza liniarităţii relaţiilor matematice, e_3 poate fi distinct de variabilele x_1 şi x_2, astfel încît putem scrie :

$$E(e_3x_1) = E\,(e_3x_2) = 0.$$

În cazul independenţei variabilelor x_1 şi x_2, putem scrie ecuaţia :

$$x_2 = a_{12}x_1 + e_2,$$

$$E\,(e_2x_1) = 0.$$

În general, aşa-zisele modele „recursive", adică non-independente, pot fi formulate astfel [8] :

$$x_n = a_{1n}\ x_1 + a_{2n}\ x_2 + \ldots + a_{(n-1)}\ x_{n-1} + e_n\ ,$$

$$b_{1n}\, x_1 + b_{2n}\, x_2 + \ldots + b_{(n-1)n}\, x_{n-1} + x_n = e_n\ ,$$

$$E\,(e_nx_j) = 0 \quad (j < n).$$

Scanned with OKEN Scanner

Fiecare variabilă explicită este funcţie liniară de variabilele explicite de rang inferior, precum şi de o variabilă implicită de acelaşi rang, independentă de aceste variabile.

Sistemul de ecuaţii cuprinzînd $\frac{n(n-1)}{2}$ necunoscute (coeficienţi) şi $\frac{n(n-1)}{2}$ ecuaţii poate fi obţinut prin formularea următoarelor ipoteze simplificatoare [4]:

$$E(e_n\, x_j) = 0,$$

$$b_{1n}\; E(x_1\, x_j) + b_{2n}\; E\,(x_2 x_j) + \ldots + E(x_n\, x_j) = 0,$$

reîncepînd pentru e_{n-1} etc.

Un astfel de sistem de ecuaţii poate constitui un model *recursiv complet*, dacă coeficienţii b_{ij} sînt apriori susceptibili de a lua o valoare diferită de zero. S i m o n a considerat ordinea stabilită între variabile drept „ordinea cauzală”[4], care este în acelaşi timp o ordine între formele liniare, stabilind relaţii între aceste variabile. Această ordine poate fi totală dacă coeficienţii variabilelor (b_{ji}) sînt, pe fiecare linie, diferiţi de 0 [8].

Matricea coeficienţilor de dependenţă b_{ji} este o matrice triunghiulară de forma matricii 14.

$$\begin{pmatrix} 1 & 0 & 0 \\ b_{12} & 1 & 0 \\ b_{13} & b_{23} & 1 \end{pmatrix}$$

Matricea 14. — Matricea coeficienţilor de dependenţă b_{ji}

Într-un model recursiv, termenii din dreapta diagonalei sînt întotdeauna nuli, neavînd nici o semnificaţie deosebită. În schimb, nulitatea unuia din termenii situaţi la stînga

[4] „*causal ordering*”.

diagonalei indică absența unei relații cauzale de ordinul $m - 1$, unde m este rangul liniei.

Atunci cînd formulăm ipoteza independenței variabilelor implicite în raport cu variabilele de rang inferior apărute pe aceiași linie, putem afirma că variabilele implicite e, luate două cîte două, sînt independente. Această ipoteză numită „disociabilitate statistică a cauzelor" poate fi dedusă, cum a arătat B o u d o n, din ipoteza „specificității erorilor", extinsă chiar la „specificitatea variabilelor implicite". Ultima ipoteză corespunde ideii de „închidere" a sistemului cauzal; fără ea, presupunem noi ecuații de tipul: $e_2 = \eta_{12}\, e_1$ [2].

Ipoteza disociabilității statistice a cauzelor rămîne aplicabilă în toate cazurile unui sistem cauzal neînchis. De pildă ansamblul de ecuații [8]:

$$b_{12}x_1 + x_2 = e_2,$$

$$b_{13}x_1 + b_{23}\, x_2 + x_3 = e_3,$$

$$E\,(e_2x_1) = 0,\ E\,(e_3x_1) = 0,\ E\,(e_3x_2) = 0,$$

care apare ca:

$$x_3 = a_{13}x_1 + a_{23}x_2 + e_3,$$

poate fi rezolvat scriind ecuațiile:

$$b_{12}\,E(x_1^2) + b_{23}\,E\,(x_1x_2) + E\,(x_1x_3) = E\,(e_2x_1) = 0;$$

$$b_{13}\,E(x_1x_2) + b_{23}\,E(x_2^2) + E\,(x_2x_3) = E\,(e_3\,x_2) = 0,$$

Aceste ecuații permit să se calculeze a_{13} și a_{23}, apoi:

$$b_{12}\,E(x_1^2) + E\,(x_1x_2) = F\,(e_2\,x_1) = 0.$$

Se observă că această ultimă ecuaţie se rezolvă prin determinarea coeficienţilor de dependenţă [8]:

$$b_{12} = -\frac{E(x_1^2 x_2)}{E(x_1^2)}$$

Deoarece $b = -a$, putem scrie:

$$a_{12} = \frac{E(x_1 x_2)}{E(x_1^2)}$$

Aici a_{12} este coeficientul de regresie al lui x_2 în raport cu x_1.

Valorile lui a_{12} se determină utilizînd calculul regresiei multiple clasice [19].

4.7. CALCULUL REGRESIEI LINIARE MULTIPLE

În calculul regresiei liniare multiple se utilizează matricea seriilor factoriale ale variabilelor aleatoare $x_1, x_2, \ldots, x_n$ şi y, reprezentînd datele statistice sau informaţiile obţinute din observaţii, de forma acelora din următoarea tabelă:

y	x_1	x_2	x_3	...	x_n
y_1	x_{11}	x_{12}	x_{13}	...	x_{1n}
y_2	x_{21}	x_{22}	x_{23}	...	x_{2n}
y_3	x_{31}	x_{32}	x_{33}	...	x_{3n}
⋮	⋮	⋮	⋮	⋮ ⋮ ⋮	⋮
y_m	x_{m1}	x_{m2}	x_{m3}	...	x_{mn}

Fiind dată această matrice, se cere să se determine coeficienţii regresiei liniare multiple de forma:

$$y = a_1x_1 + a_2x_2 + \ldots + a_nx_n + e_y.$$

Coeficienţii a_j, unde $j = 0, 1, 2, 3, \ldots, n$, se determină din condiţiile:

$$y_i = e_y + \sum_{j=1}^{n} a_j \cdot x_{ij}, \quad i = 1, 2, 3, \ldots, m.$$

Deoarece $m \neq n + 1$, acest sistem de ecuaţii liniare este nedeterminat. De aceea, în scopul ridicării nedeterminării se poate utiliza metoda celor mai mici pătrate, făcînd ipoteza că dispersia $D(y)$ este minimă, adică:

$$D(y) = \frac{1}{m} \sum_{i=1}^{m} \left(y_i - e_y - \sum_{j=1}^{n} a_i\, x_{ij} \right)^2 = \min.$$

În acest caz, coeficienţii a_j, $j = 0, 1, 2, \ldots, n$ se determină din condiţiile:

$$\frac{\partial D(y)}{\partial e_y} = 0, \quad \frac{\partial D(y)}{\partial x_j} = 0, \; j = 1, 2, \ldots, n,$$

$$\frac{\partial D(y)}{\partial e_y} = 0, \quad \sum_{i=1}^{m} \left(y_i - e_y - \sum_{j=1}^{n} a_j\, x_{ij} \right)(-1) = 0.$$

Din aceste condiţii rezultă:

a) pentru $\dfrac{\partial D(y)}{\partial e_y} = 0$:

$$\sum_{i=1}^{m} y_i - \sum_{i=1}^{m} e_y - \sum_{i=1}^{m} \sum_{j=1}^{m} a_j x_{ij} = 0,$$

de unde:

$$me_y + \sum_{j=1}^{n}\left(\sum_{i=1}^{m} x_{ij}\right) e_y = \left(\sum_{i=1}^{m} y_i\right).$$

b) pentru $\frac{\partial D(y)}{\partial e_y} = 0$:

$$2\sum_{i=1}^{m}\left(y_i - e_y - \sum_{j=1}^{n} a_j x_{ij}\right)(-x_{ik}) = 0, \quad k = 1, 2, 3, \ldots, n,$$

de unde:

$$\left(\sum_{i=1}^{m} y_i\right) e_y + \sum_{j=1}^{n}\left(\sum_{i=1}^{n} x_{ij}x_{ik}\right) a_j = \left(\sum_{i=1}^{m} y_i x_{ik}\right).$$

În felul acesta se obţine sistemul cu $n+1$ ecuaţii avînd $n+1$ necunoscute $a_1, a_2, \ldots, a$ şi e_y;

$$me_y + \sum_{j=1}^{n}\left(\sum_{i=1}^{m} x_{ij}\right) a_j = \left(\sum_{i=1}^{m} y_i\right)$$

$$\left(\sum_{i=1}^{m} x_{ik}\right) e_y + \sum_{j=1}^{n}\left(\sum_{i=1}^{m} x_{ij}x_{ik}\right) a_j = \sum_{i=1}^{m} (y_i x_{ik}).$$

Rezolvarea acestui sistem micşorează precizia calculelor, întrucît seriile factoriale ale variabilelor aleatoare $x_1, x_2, \ldots, x_n, y$ pot avea ordini de mărimi foarte diferite, şi orice normalizare a lor poate produce erori însemnate. Pentru a fi înlăturat acest inconvenient, sistemul de mai sus se va transforma astfel încît, în locul seriilor factoriale se vor introduce indicatorii media aritmetică şi abaterea medie pătratică corespunzătoare acestor serii factoriale.

Scanned with OKEN Scanner

Ţinînd seama că :

$$\sum_{i=1}^{m} x_{ij} = m \cdot \overline{x}_j, \quad j = 1, 2, \ldots, n$$

$$\sum_{i=1}^{m} y_i = m \cdot \overline{y}.$$

Înlocuind aceste relaţii în sistemul de ecuaţii precedent, obţinem :

$$me_y + \sum_{j=1}^{n} m\,\overline{x}_j a_j = m\,\overline{y}$$

$$(m\,\overline{x}_k)\,e_y + \sum_{j=1}^{n} \left(\sum_{i=1}^{m} x_{ij} x_{ik} \right) a_j = \left(\sum_{i=1}^{m} y_i x_{ik} \right).$$

Simplificînd aceste ecuaţii, obţinem :

$$e_y + \sum_{j=1}^{n} \overline{x}_j a_j = \overline{y}$$

$$(m\,\overline{x}_k)\,e_y + \sum_{j=1}^{n} \left(\sum_{i=1}^{m} x_{ij} x_{ik} \right) a_j = \left(\sum_{i=1}^{m} y_i x_{ik} \right).$$

Eliminînd pe e_y, obţinem :

$$m\,\overline{x}_k \left(\overline{y} - \sum_{j=1}^{n} \overline{x}_j a_j \right) + \sum_{j=1}^{n} \left(\sum_{i=1}^{n} x_{ij} x_{ik} \right) a_j = \sum_{i=1}^{m} y_i x_{ik}.$$

Prin urmare :

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{m} x_{ij} x_{ik} - m\,\overline{x}_j\,\overline{x}_k \right) a_j = \left(\sum_{i=1}^{m} y_i x_{ik} \right) - m\,\overline{y}\,\overline{x}_k.$$

Scanned with OKEN Scanner

Dacă luăm în considerație și al aterile medii pătratice σ_j, σ_k, σ_y și efectuînd unele artificii de calcul, se obține :

$$\frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{m} x_{ij}x_{ik} - m\,\bar{x}_j\bar{x}_k\right)}{m\,\sigma_j\,\sigma_k}\,\sigma_j a_j = \frac{\left(\sum_{i=1}^{m} y_i x_{ik} - m\,\bar{y}\,\bar{x}_k\right)}{m\,\sigma_y\,\sigma_k}\cdot\sigma_y$$

Cum însă

$$\frac{\sum_{i=1}^{m} x_{ij}x_{ik} - m\,\bar{x}_j\bar{x}_k}{m\,\sigma_j\,\sigma_k} = r_{jk}, \quad \frac{\sum_{i=1}^{m} y_i x_{ik} - m\,\bar{y}\,\bar{x}_k}{m\,\sigma_y\,\sigma_k} = r_{yk},$$

sînt coeficienții de corelație între perechile de variabile aleatoare (x_j, x_k) respectiv (y, x_k) și făcînd schimbarea de variabilă, putem scrie :

$$\sum_{j=1}^{n} r_{jk}b_j = r_{yk}, \quad k = 1, 2, \ldots, n.$$

Sub această formă, sistemul de ecuații este ușor de rezolvat, întrucît coeficienții r_{jk} sînt numere relativ subunitare ($|r_{jk}| \leqslant 1$) și în plus este îndeplinită și condiția de obținere a preciziei maxime prin metode exacte, care constă în aceea că elementele diagonalei principale să fie maxime pe linia și coloana care le conțin ($r_{jj} = 1$).

Ținînd seama de seriile factoriale ale variabilelor aleatoare independente $x_1, x_2, \ldots, x_n$ și respectiv ale variabilei aleatoare dependente y, pot fi determinați coeficienții de regresie liniară a_j, $j = 0, 1, 2, \ldots, n$; în funcție de valoarea coeficientului multiplu de corelație se poate stabili și intensitatea influenței liniare a variabilelor independente (x_j) asupra variabilei dependente (y).

Fiind dată matricea seriilor factoriale, se calculează pe rînd :

a) mediile aritmetice $\overline{x}_j$ și $\overline{y}$ (tabela de mai jos):

$$\overline{x}_j = \frac{1}{m} \sum_{i=1}^{m} x_{ij}, \quad j = 1, 2, \ldots, n, n+1.$$

$$\overline{x}_{n+1} = \overline{y}.$$

Tabela mediilor aritmetice ale seriilor factoriale ale variabilelor aleatoare este următoarea:

y	x_1	x_2	x_3	$\ldots$	x_n
y_1	x_{11}	x_{12}	x_{13}	$\ldots$	x_{1n}
y_2	x_{21}	x_{22}	x_{23}	$\ldots$	x_{2n}
y_3	x_{31}	x_{32}	x_{33}	$\ldots$	x_{3n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_m	x_{m1}	x_{m2}	x_{m3}	$\ldots$	x_{mn}
$\overline{y}$	$\overline{x}_1$	$\overline{x}_2$	$\overline{x}_3$	$\ldots$	$\overline{x}_n$

b) abaterile medii pătratice σ_j:

$$\sigma_j = \frac{\sum_{i=1}^{m} x_{ij}^2 - \frac{1}{m}\left(\sum_{i=1}^{m} x_{ij}\right)^2}{n} \quad j = 1, 2, \ldots, n, n+1,$$

$$\sigma_{n+1} = \sigma_y.$$

c) coeficienții de corelație r_{jk} (tabela de mai jos):

$$r_{jk} = \frac{\frac{1}{m}\sum_{i=1}^{m} x_{ij}x_{ik} - \overline{x}_j\overline{x}_k}{\sigma_j \sigma_k}, \quad \text{unde: } r_{k,n+1} = r_{ky} = r_{yk},$$

$$k = 1, 2, 3, \ldots, n,$$
$$j = 1, 2, 3, \ldots, n, n+1.$$

Valorile coeficienților de corelație vor fi :

$$r_{jk} = r_{kj},$$
$$|r_{kj}| < 1, \quad k \neq j,$$
$$r_{jj} = 1.$$

Tabela coeficienților de corelație statistică este următoarea:

σ_y	σ_1	σ_2	σ_3	...	σ_n
r_{1y}	1	r_{12}	r_{13}	...	r_{1n}
r_{2y}	r_{21}	1	r_{23}	...	r_{2n}
r_{3y}	r_{31}	r_{32}	1	...	r_{3n}
⋮	⋮	⋮	⋮	⋮	⋮
r_{ny}	r_{n1}	r_{n2}	r_{n3}	...	1

d) sistemul ecuațiilor normate (standard) de regresie :

$$\sum_{j=1}^{n} r_{kj} b_j = r_{ky}, \quad k = 1, 2, \ldots, n,$$

$$b_x = (r_{kj})_{kj}, \quad b = (b_j)_j, \quad r_y = (r_{ky})_k,$$

$$r_x b = r_y, \quad b = r_x^{-1} \cdot r_y = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

e) coeficienții regresiei liniare :

$$a_j = b_j \cdot \frac{\sigma_y}{\sigma_j}, \quad j = 1, 2, \ldots, n.$$

$$e_y = \overline{y} - \sum_{j=1}^{n} a_j \cdot x_j.$$

f) ecuaţia de regresie liniară [20]:

$$y = a_1x_1 + a_2x_2 + \ldots + a_mx_n + e_y.$$

În această ecuaţie, coeficienţii a_j şi valoarea (eroarea) reziduală e_y sînt cunoscute. Prin înlocuirea în ecuaţia de regresie a valorilor concrete ale variabilelor independente x_j, obţinem valoarea probabilă a variabilei dependente y. Stabilind această dependenţă dintre variabilele structurii cauzale complexe, am formulat astfel un model matematic de analiză cauzală (de dependenţă) [21].

4.8. UN EXEMPLU DE ANALIZĂ A RELAŢIILOR DE DEPENDENŢĂ ÎNTRE NIVELUL DE INSTRUIRE, GRADUL DE CALIFICARE ŞI PRODUCTIVITATEA MUNCII

Un ansamblu de coeficienţi de dependenţă poate fi asociat unei structuri cauzale identificabile. De pildă, într-o astfel de structură, x_i este dependent de x_j, dacă speranţa matematică este $E(e_ix_j) = 0$, putem scrie ecuaţia de dependenţă în forma ei generală [8].

$$b_{1i}E(x_1x_j) + \ldots + E(x_ix_j) + \ldots + b_{ji}E(x_j^2) + \ldots + b_{mi}E(x_mx_j) = E(e_ix_j) = 0.$$

În cazul unei structuri identificabile, dacă admitem ipoteza de noncorelaţie între factorii impliciţi, putem utiliza un mecanism de rezolvare a sistemului de ecuaţii de forma celor de la 4.3. Cu ocazia rezolvării acestor sisteme, S i m o n a atribuit coeficienţilor o interpretare cauzală. În fapt, coeficienţii de dependenţă obţinuţi au reprezentat o extindere a coeficienţilor de regresie [8]. Mai mult chiar, în cazul unor structuri cauzale identificabile, coeficienţii lui Wright şi coeficienţii de degependenţă standardizaţi sînt o extindere a coeficienţilor de corelaţie parţiali.

Scanned with OKEN Scanner

În general, în studiul inferenţelor cauzale existente în structurile cauzale analizate este indiferent faptul dacă utilizăm un algoritm pentru analiza regresiei sau pentru analiza dependenţei. Totuşi, dacă dorim să utilizăm analiza de dependenţă, trebuie să apelăm ori la coeficientul de dependenţă (b_{ji}) ori la coeficientul lui Wright $\left(b_{ji}\frac{\sigma_j}{\sigma_i}\right)$. De cele mai multe ori însă, pentru a facilita interpretarea şi efectuarea calculelor, utilizarea coeficientului lui Wright prezintă avantajul exprimării exclusive a acestuia în funcţie de coeficienţii de corelaţie [7], [8]. Ecuaţiile în care intervin coeficienţii de dependenţă (b_{ji}), conţin dispersiile factorilor expliciţi. Revenind la ecuaţia scrisă mai înainte, în care înlocuim pe $E(x_1x_j)$ cu $\sigma_1\sigma_j r_{1j}$, ..., $E(x_ix_j)$ cu $\sigma_i\sigma_j r_{ij}$ etc. obţinem [4]

$$b_{1i}\sigma_1\sigma_j r_{1j}+\ldots+\sigma_i\sigma_j r_{ij}+\ldots+b_{ji}\sigma_j^2+\ldots+b_{mi}\sigma_m\sigma_j r_{mi}=0.$$

Această ecuaţie nu se modifică în mod evident atunci cînd multiplicăm fiecare monom cu $\frac{\sigma_i}{\sigma_i}=1$:

$$b_{1i}\frac{\sigma_1}{\sigma_i}\sigma_i\sigma_j r_{1j}+\ldots+\sigma_i\sigma_j r_{ij}+\ldots+b_{ji}\frac{\sigma_j}{\sigma_i}\sigma_i\sigma_j+$$

$$+\ldots+b_{mi}\frac{\sigma_m}{\sigma_i}\sigma_i\sigma_j r_{mj}=0.$$

Eliminînd factorii $\sigma_i\sigma_j$ din ecuaţie şi remarcînd cantităţile $b_{1i}\frac{\sigma_1}{\sigma_i}$, ..., $b_{ji}\frac{\sigma_j}{\sigma_i}$, ..., $b_{mi}\frac{\sigma_m}{\sigma_i}$, care sînt coeficienţii lui Wright definiţi prin $d_{1i}=\frac{\sigma_{ie}}{\sigma_i}\frac{\sigma_1}{\sigma_{1e}}=a_{1i}\frac{\sigma_1^2}{\sigma_i^2}$, putem scrie:

$$d_{1i}r_{1j}+\ldots+r_{ij}+\ldots+d_{ji}+\ldots+d_{mi}r_{mj}=0.$$

Inv. 117649

Scanned with OKEN Scanner

Examinînd această ecuaţie, observăm că am obţinut coeficienţii lui Wright în funcţie doar de coeficienţii de corelaţie [8]. De aceea, analiza matematică se simplifică foarte mult, deoarece este suficient să calculăm toţi coeficienţii de corelaţie simplă între variabilele explicite pentru a obţine coeficienţii lui Wright, cu ajutorul cărora putem măsura dependenţa cauzală.

Vom da un exemplu de aplicare a analizei de dependenţă pentru modelul din figura 40.

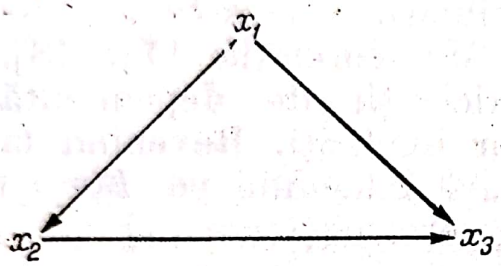


Fig. 40. — Structura relaţiilor dintre nivelul de şcolarizare (x_1), gradul de calificare (x_2) şi productivitatea muncii (x_3)

Ecuaţiile cauzale pentru modelul din figura 40 sînt:

$$x_1 = e_1,$$
$$b_{12}x_1 + x_2 = e_2,$$
$$b_{13}x_1 + b_{23}x_2 + x_3 = e_3.$$

Ipotezele restrictive $E\ (e_i\ x_j)$ se aplică astfel:

$$E(e_1e_2) = E(e_1e_3) = E(e_2x_1) = E(e_3x_1) = E(e_1x_2) =$$
$$= E(e_2x_2) = E(e_3x_2) = E(e_1x_3) = E(e_2x_3) = E(e_3x_3) = 0.$$

Făcînd unele ipoteze restrictive de acest fel putem obţine următoarele ecuaţii de dependenţă:

$$b_{12}\,E(x_1x_3) + E(x_2x_3) = 0,$$
$$b_{13}\,E(x_1^2) + b_{23}\,E(x_1x_2) + E(x_1x_3) = 0,$$
$$b_{13}\,E(x_1x_2) + b_{23}\,E(x_2^2) + E(x_2x_3) = 0.$$

Scriind speranţele matematice $E(x_ix_j)$ sub forma:

$$E(x_1x_2) = \sigma_1\,\sigma_2 r_{12},\quad E(x_1x_3) = \sigma_1\,\sigma_3 r_{13},$$
$$E(x_2x_3) = \sigma_2\,\sigma_3 r_{23},\quad E(x_1^2) = \sigma_1^2,\quad E(x_2^2) = \sigma_2^2$$ şi efectuînd

înlocuirile în ecuaţiile precedente, obţinem :

$$b_{12}\sigma_1\sigma_3 r_{13} + \sigma_2\sigma_3 r_{23} = 0,$$
$$b_{13}\sigma_1^2 + b_{23}\sigma_1\sigma_2 r_{12} + \sigma_1\sigma_3 r_{13} = 0,$$
$$b_{13}\sigma_1\sigma_2 r_{12} + b_{23}\sigma_2^2 + \sigma_2\sigma_3 r_{23} = 0.$$

Aceste ecuaţii pot fi modificate prin multiplicarea fiecărui monom cu $\frac{\sigma_1}{\sigma_1} = \frac{\sigma_2}{\sigma_2} = \frac{\sigma_3}{\sigma_3} = 1$, astfel încît :

$$b_{12}\frac{\sigma_1}{\sigma_2^2}\sigma_2\sigma_3 r_{13} + \sigma_2\sigma_3 r_{23} = 0,$$

$$b_{13}\frac{\sigma_1}{\sigma_3}\sigma_1\sigma_3 + b_{23}\frac{\sigma_2}{\sigma_3}\sigma_1\sigma_3 r_{12} + \sigma_1\sigma_3 r_{13} = 0,$$

$$b_{13}\frac{\sigma_1}{\sigma_3}\sigma_2\sigma_3 r_{12} + b_{23}\frac{\sigma_2}{\sigma_3}\sigma_2\sigma_3 + \sigma_2\sigma_3 r_{13} = 0.$$

Simplificînd ecuaţiile cu $\sigma_2\sigma_3$ şi $\sigma_1\sigma_3$ şi remarcînd că $d_{12} = b_{12}\frac{\sigma_1}{\sigma_2}$, $d_{13} = b_{13}\frac{\sigma_1}{\sigma_2}$ şi $d_{23} = b_{23}\frac{\sigma_2}{\sigma_3}$ sînt coeficienţii lui Wright, obţinem sistemul :

$$d_{12}r_{13} + r_{23} = 0,$$
$$d_{13} + d_{23}r_{12} + r_{13} = 0,$$
$$d_{13}r_{12} + d_{23} + r_{23} = 0.$$

Prin urmare, am obţinut o serie de relaţii între coeficienţii de dependenţă şi coeficienţii de corelaţie. Aplicînd în aceste relaţii diverşi coeficienţi de corelaţie observaţi, putem scrie :

$$r_{12} = 0{,}80,\ r_{23} = 0{,}72,\ r_{23} = 0{,}50.$$

Scanned with OKEN Scanner

Înlocuind aceşti coeficienţi în relaţiile precedente, obţinem soluţiile sistemului :

$$d_{12} = -0{,}64\,; \quad d_{13} = -0{,}54\,; \quad d_{23} = -0{,}22.$$

Cantităţile d_{12}, d_{13} şi d_{23} sînt respectiv egale cu $-d_{12}$, $-d_{13}$ şi $-d_{23}$, astfel încît : $d_{12} = 0{,}64$, $d_{13} = 0{,}54$, $d_{23} = 0{,}22$.

Schimbarea de semn a coeficienţilor lui Wright a fost necesară întrucît ecuaţiile cauzale sînt scrise sub forma :

$b_{12}x_1 + \ldots + e_i = e_i$, şi nu sub forma $x_1 = a_{12}x_1 + \ldots + e_i$.

Examinînd parametrii d_{ji} din ecuaţiile de dependenţă cauzală, am obţinut valorile teoretice, ajustate, ale coeficienţilor de corelaţie, pe care le vom compara cu valorile lor observate, astfel încît :

	Valori observate	Valori teoretice
r_{12}	0,80	0,52
r_{13}	0,72	0,72
r_{23}	0,50	0,50

Comparînd valorile teoretice cu cele observate ale coeficienţilor de corelaţie putem aprecia calitatea ajustării obţinute.

4.9. UN ALT EXEMPLU : MODELUL CAUZAL AL PROCESULUI DE URBANIZARE

În studiul structurilor sociale concrete pot fi utilizate două metode matematice de analiză a indicatorilor deductivi : *analiza corelaţională* şi *analiza cauzală* [4]. Este evi-

dent că analiza cauzală este superioară analizei corelaţionale. Cu toate acestea, efectuarea unei analize cauzale fără ca în prealabil să se efectueze o analiză corelaţională este extrem de dificilă [23].

Greutăţile cele mai mari în stabilirea unei legături între o măsură a corelaţiei şi o interpretare cauzală sînt generate de dificultatea eliminării efectelor variabilelor care pot produce o relaţie cauzală aparentă, falsă. Aceste greutăţi pot fi însă învinse prin construirea unui *model cauzal* care să permită deducerea relaţiilor cauzale din corelaţiile statistice, ceea ce este pe deplin posibil [4].

Întrucît corelaţiile statistice participă la realizarea inferenţelor cauzale, relaţiile de dependenţă cauzală pot fi deduse din *modelul corelaţiilor* [24].

Spre exemplificare, ne vom baza pe modelul corelaţiilor statistice din figura 41. Acest model reprezintă de fapt un graf neorientat (fără arce) al corelaţiilor existente între patru variabile (caracteristici);

— producţia industrială (X_1);
— numărul de muncitori (X_2);
— acţiuni-social-culturale (X_3);
— gradul de urbanizare (Y).

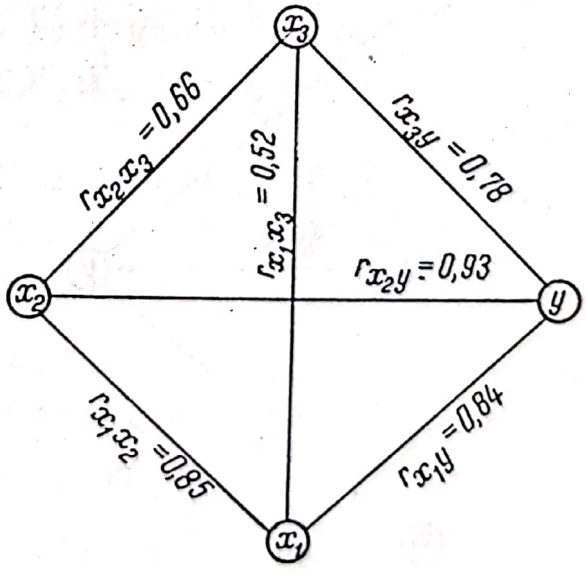


Fig. 41 — Graful neorientat al corelaţiilor statistice

Modelului matematic din figura 40 i se poate asocia matricea 15.

	x_1	x_2	x_3	y
x_1		0,85	0,52	0,84
x_2	0,85		0,66	0,93
x_3	0,52	0,66		0,78
y	0,84	0,93	0,78	

Matricea 15. — Matricea de corelaţie asociată grafului neorientat din fig. 41

Emiţînd ipoteza că producţia industrială (x_1), numărul de muncitori (x_2) şi acţiunile social-culturale (x_3) formează ansamblul cauzal al gradului de urbanizare (y) construim modelul cauzal din figura 42. Acest model reprezintă un graf orientat al relaţiilor cauzale între un x particular (x_j) şi y.

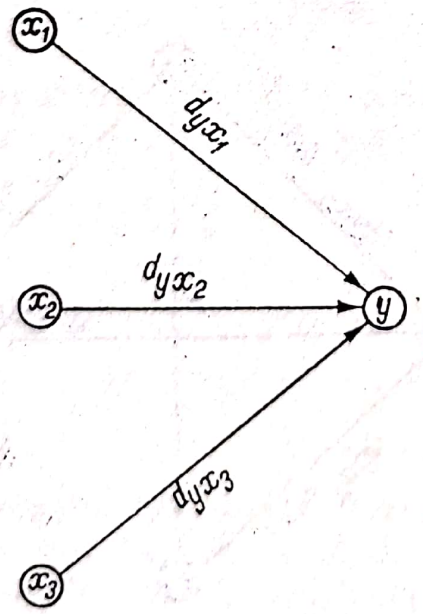


Fig. 42. — Graful orientat al relaţiilor cauzale

Ca să verificăm ipoteza emisă asupra ansamblului cauzal (x_i) al lui y vom asocia grafului orientat din figura 42 o ecuaţie de regresie liniară de forma :

$$Y = a_{x_1 y} X_1 + a_{x_2 y} X_2 + a_{x_3 y} X_3 + e_y,$$

unde :

$Y = y - \bar{y}$;
$X_j = x_j - \bar{x}_j$;
$a_{x_j y}$ — coeficienţii de regresie liniară ;
e_y — eroare reziduală.

Multiplicînd ecuaţia de regresie în mod succesiv cu X_1, X_2, şi X_3, după care însumînd fiecare monom în parte, obţinem sistemul de ecuaţii liniare :

$$\Sigma X_1 Y = a_{x_1 y} \Sigma X_1^2 + a_{x_2 y} + \Sigma X_1 X_2 + a_{x_3 y} \Sigma X_1 X_3 + \Sigma e_x X_1,$$
$$\Sigma X_2 Y = a_{x_1 y} \Sigma X_1 X_2 + a_{x_2 y} \Sigma X_2^2 + a_{x_3 y} \Sigma X_2 X_3 + \Sigma e_y X_2.$$
$$\Sigma X_3 Y = a_{x_1 y} \Sigma X_1 X_3 + a_{x_2 y} \Sigma X_2 X_3 + a_{x_3 y} \Sigma X_3^2 + \Sigma e_y X_3.$$

Făcînd ipoteza că erorile reziduale e_y sînt necorelate cu variabilele X_j, vom avea $\Sigma e_y X_1 = \Sigma e_y X_2 = \Sigma e_y X_3 = 0$, iar sistemul de ecuaţii liniare se va putea scrie astfel :

$$\Sigma X_1 Y = a_{x_1 y} \Sigma X_1^2 + a_{x_2 y} \Sigma X_1 X_2 + a_{x_3 y} \Sigma X_1 X_3,$$
$$\Sigma X_2 Y = a_{x_1 y} \Sigma X_1 X_2 + a_{x_2 y} \Sigma X_2^2 + a_{x_3 y} \Sigma X_2 X_3,$$
$$\Sigma X_3 Y = a_{x_1 y} \Sigma X_1 X_3 + a_{x_2 y} \Sigma X_2 X_3 + a_{x_3 y} \Sigma X_3^2.$$

În acest sistem de ecuaţii putem pune în evidenţă o serie de speranţe matematice :

$$\Sigma X_1 Y = N \sigma \dot{x}_1 \sigma_x r_{x_1 y} ; \Sigma X_2 Y = N \sigma_{x_2} \sigma_y r_{x_2 y} ; \Sigma X_3 Y = N \sigma_{x_3} \sigma_x r_{x_3 y}$$
$$\Sigma X_1 X_2 = N \sigma_{x_1} \sigma_{x_2} r_{x_1 x_2} ; \Sigma X_1 X_3 = N \sigma_{x_1} \sigma_{x_3} r_{x_1 x_3} ;$$
$$\Sigma X_2 X_3 = N \sigma_{x_2} \sigma_{x_3} r_{x_2 x_3} ;$$
$$\Sigma X_1^2 = N \sigma_{x_1^2} ; \Sigma X_2^2 = N \sigma_{x_2^2} ; \Sigma X_3^2 = N \sigma_{x_3^2}.$$

Dacă înlocuim speranţele matematice în ecuaţiile precedente, obţinem :

$$\sigma_{x_1}\ \sigma_y\, r_{x_1 y} = a_{x_1 y}\sigma_{x_1}^2 + a_{x_2 y}\sigma_{x_1}\ \sigma_{x_2}\ r_{x_1 x_2} + a_{x_3 y}\ \sigma_{x_1}\ \sigma_{x_3} r_{x_1 x_3},$$

$$\sigma_{x_2}\ \sigma_y r_{x_2 y} = a_{x_1 y}\ \sigma_{x_1}\ \sigma_{x_2} r_{x_1 x_2} + a_{x_2 y}\ \sigma_{x_2}^2 + a_{x_3 y}\sigma\, x_2\ \sigma\, x_3 r_{x_2 x_3}.$$

$$\sigma_{x_3}\ \sigma_y\ r_{x_3 y} = a_{x_1 y}\ \sigma_{x_1}\ \sigma_{x_3}\ r_{x_1 x_3} + a_{x_2 y}\sigma_{x_2}\ \sigma_{x_3}\ r_{x_2 x_3} + a_{x_3 y}\sigma_{x_3}^2.$$

Aceste ecuaţii pot fi modificate prin multiplicarea fiecărui monom care conţine pe $a_{x_j y}$ cu raportul $\frac{\sigma_y}{\sigma_y}$, astfel încît :

$$r_{x_1 y} = a_{x_1 y}\frac{\sigma_{x_1}}{\sigma_y} + r_{x_1 x_2}\ a_{x_2 y}\frac{\sigma_{x_2}}{\sigma_y} + r_{x_1 x_3}\ a_{x_3 y}\frac{\sigma_{x_3}}{\sigma_y},$$

$$r_{x_2 y} = r_{x_1 x_2}\ a_{x_1 y}\frac{\sigma_{x_1}}{\sigma_y} + a_{x_2 y}\frac{\sigma_{x_2}}{\sigma_y} + r_{x_2 x_3}\ a_{x_3 y}\frac{\sigma_{x_1}}{\sigma_y}.$$

$$r_{x_3 y} = r_{x_1 x_3}\ a_{x_1 y}\frac{\sigma\, x_1}{\sigma\, y} + r_{x_2 x_3}\ a_{x_2 y}\frac{\sigma\, x_2}{\sigma\, y} + a_{x_3 y}\frac{\sigma_{x_2}}{\sigma_y}.$$

Remarcînd că $d_{x_j y} = a_{x_j v}\frac{\sigma_{x_1}}{\sigma_y}$ este coeficientul lui Wright, sistemul de ecuaţii poate fi scris astfel :

$$r_{x_1 y} = d_{x_1 y} + r_{x_1 x_2}\ d_{x_2 y} + r_{x_1 x_3}\ d_{x_3 y}.$$

$$r_{x_2 y} = r_{x_1 x_2}\ d_{x_1 y} + d_{x_2 y} + r_{x_2 x_3}\ d_{x_3 y}.$$

$$r_{x_3 y} = r_{x_1 x_3}\ d_{x_1 y} + r_{x_2 x_3}\ d_{x_2 y} + d_{x_3 y}.$$

Scanned with OKEN Scanner

Înlocuind valorile coeficienților de corelație din matricea de corelație (matricea 12), obținem sistemul de ecuații:

$$0{,}84 = d_{x_1 y} + 0{,}85\ d_{x_2 y} + 0{,}52\ d_{x_3 y}.$$

$$0{,}93 = 0{,}85\ d_{x_1 y} + d_{x_2 y} + 0{,}66\ d_{x_3 y}.$$

$$0{,}78 = 0{,}52\ d_{x_1 y} + 0{,}66\ d_{x_2 y} + d_{x_3 y}.$$

Numărul necunoscutelor $d_{x_j y}$ fiind egal cu numărul ecuațiilor, sistemul de ecuații liniare poate fi rezolvat cu ajutorul formulelor lui Cramer [25].

$$d_{x_j y} = \frac{D_j}{|D|},\ j = 1, 2, 3, \text{ în care:}$$

$|D|$ — determinantul matricei sistemului de ecuații dat;
D_j — determinantul obținut din $|D|$ prin înlocuirea coloanei j cu termenii liberi.

Aplicînd formulele lui Cramer s-au obținut valorile determinanților:

$$|D| = 0{,}155;\ D_1 = 0{,}035;\ D_2 = 0{,}083;\ D_3 = 0{,}046.$$

Pe baza valorilor determinanților D_j și $|D|$ s-au putut calcula coeficienții de dependență ($d_{x_j y}$):

$$d_{x_1 y} = \frac{D_1}{|D|} = \frac{0{,}035}{0{,}155} = 0{,}225;\ d_{x_2 y} = \frac{D_2}{|D|} = \frac{0{,}083}{0{,}155} = 0{,}535;$$

$$d_{x_3 y} = \frac{D_3}{|D|} = \frac{0{,}046}{0{,}155} = 0{,}296.$$

Scanned with OKEN Scanner

O succintă analiză a coeficienţilor de dependenţă ($d_{x_j y}$) ne evidenţiază faptul că gradul de urbanizare (y) este dependent de producţia industrială (x_1), de numărul de muncitori (x_2) şi de acţiunile social-culturale (x_3).

Intensitatea relaţiilor cauzale semnalate este însă diferită, deoarece dependenţa gradului de urbanizare (y) de numărul de muncitori (x_2) este de intensitate medie iar producţia industrială (x_1) şi acţiunile social-culturale (x_3) influenţează relativ slab gradul de urbanizare (y)[5].

Evidenţiind o dependenţă cauzală a gradului de urbanizare (y) de producţia industrială (x_1), de numărul de muncitori (x_2) şi de acţiunile social culturale (x_3) am verificat justeţea ipotezei emise asupra relaţiilor cauzale din structura socială analizată.

Comparînd coeficienţii de corelaţie ($r_{x_j y}$) cu coeficienţi de dependenţă cauzală ($d_{x_j y}$) obţinem tabelul 17.

Tabelul 17

Coeficientul de corelaţie ($r_{x_j y}$)		Coeficientul de dependenţă ($d_{x_j y}$)	
Simbol	Valoarea coeficientului	Simbol	Valoarea coeficientului
$r_{x_1 y}$	0,840	$d_{x_1 y}$	0,225
$r_{x_2 y}$	0,930	$d_{x_2 y}$	0,535
$r_{x_3 y}$	0,780	$d_{x_3 y}$	0,296

Între coeficienţii de corelaţie ($r_{x_j y}$) şi de dependenţă ($d_{x_j y}$) din această tabelă există o deosebire evidentă. Această deosebire este explicabilă dacă ne gîndim la faptul că în timp ce o corelaţie statistică evidenţiază o asociere, un paralelism a două caracteristici explicite, o dependenţă

[5] Deoarece s-a utilizat un sistem incomplet de caracteristici (variabile) pentru caracterizarea procesului de urbanizare, concluziile trase nu sînt valabile, ci pot constitui cel mult un exemplu de analiză cauzală.

Scanned with OKEN Scanner

cauzală semnalează o relaţie statistică dintre aceste două caracteristici, din care cel puţin una este cauza unui efect exercitat asupra celeilalte.

4.10. CONCLUZII

Între noţiunile de corelaţie statistică şi de relaţie cauzală există o deosebire evidentă. Dacă prin corelaţie statistică noi vom înţelege o variaţie concomitentă a două sau mai multe variabile, printr-o relaţie cauzală vom pune în evidenţă o dependenţă între variabile, din care cel puţin una este *cauza* unui *efect* exercitat asupra celeilalte.

Întrucît corelaţiile statistice participă la realizarea inferenţelor cauzale, relaţiile de dependenţă cauzală pot fi deduse din corelaţiile statistice.

În domeniul sociologiei, structurile cauzale simple sînt foarte rare. De aceea, eliminarea efectului variabilelor care generează dependenţe aparente este extrem de dificilă.

În practică se întîlnesc relaţii numeroase, multiple, între variabile. Cercetarea doar a unor relaţii cauzale izolate din mulţimea relaţiilor existente în structura unui proces social oarecare este destul de riscantă, dacă vom avea în vedere faptul că acestea pot fi influenţate cauzal de celelalte variabile din sistem.

În condiţiile inexistenţei pluralităţii cauzelor, adică în cazul cauzei unice de tipul : „x este cauza lui y”, coeficientul de corelaţie statistică constituie o măsură bună a dependenţei cauzale dintre x şi y. Într-un caz contrar, coeficientul de corelaţie statistică reflectă cel mult o variaţie concomitentă a variabilelor corelate, neputînd da nici o indicaţie asupra inferenţelor cauzale. De aceea, în asemenea cazuri trebuie utilizat *coeficientul de dependenţă* pentru măsurarea inferenţelor cauzale, ţinînd seama de întregul ansamblu de relaţii cauzale existente în structura analizată.

Scanned with OKEN Scanner

Fenomenele şi procesele sociale implică structuri cauzale complexe, structuri care pot fi analizate eficace prin trecerea de la situaţia cauzalităţii la traducerea sa matematică. Această trecere de la sistemele de relaţii cauzale la traducerea lor matematică se poate face prin intermediul coeficienţilor de dependenţă. Utilizarea acestor coeficienţi în analiza matematică a fenomenelor şi proceselor sociale deosebit de complexe permite verificarea sistemelor de relaţii cauzale.

În prezent, metodele şi modelele de analiză cauzală stau în atenţia oamenilor de ştiinţă în vederea extinderii aplicaţiilor matematicii în ştiinţele umane, îndeosebi în domeniul sociologiei. Interesul crescînd pentru aceste metode şi modele de analiză matematică este justificat dacă avem în vedere faptul că ele pot extinde într-o măsură însemnată sfera aplicaţiilor matematicii în domeniul ştiinţelor umane. De pildă, prin perfecţionarea acestor metode şi modele matematice vor fi deschise perspective excepţionale în domeniul analizei matematice a fenomenelor şi proceselor sociale. În acest fel se vor putea construi modele matematice ale structurilor cauzale, foarte diferite, modele care vor constitui un cîştig pentru ştiinţele umaniste [26], [27].

BIBLIOGRAFIE

1. René Bassoul, *Mathematique et causalité en sociologie*, în: La Revue Française de sociologie, VIII, 1967.
2. Raymond Boudon, *A method of linear causal analysis: dependence analysis*, în: American sociological review 30, 1965.
3. Hubert M. Jr. Blalock, *Correlation and causality: the multivariate case*, în: Social forces, 30, 1961.
4. Raymond Boudon, *L'analyse mathématique des faits sociaux*, Paris, Libr, Plon, 1967.
5. R. A. Sutermeister, *People and Productivity*, McCraw-Hill.

Scanned with OKEN Scanner

6. Raymond Boudon, *Méthodes d'analyse causale.* În: Revue français de sociologie, 6, 1955.

7. Raymond Boudon şi Paul F. Lezarsfeld, *L'analyse empirique de la causalité.* La Haye-Paris, Mouton, 1966.

8. Simon Herbert A., *Spurios correlation: a causal interpretation,* în: Journal of the American Statistical Association, 49, 1954.

9. M. Vasilescu, M. Şileţchi, L. Curta şi Radu A. Almăşan *Application of Matrix Algebra to the Study of the Structure of social Processes* în: Cibernetica, 299, 1970.

10. Robinson Warren, *Asymmetric causal models: comments on Polk and Blalock,* în: American sociological review, 27, 1962.

11. Simon Herbert A. *Causality and econometrics comment,* în: Econometrica, 23, 1955.

12. T. W. Anderson, *An introduction to multivariate statistical analysis* New York, Wiley 1958.

13. Hubert M. Jr. Blalock, *From variable causal models and partial correlation,* în: The American Journal of Sociology, 68, 1962.

14. James S. Coleman, *Introduction to mathematical sociology,* Glencoe, I11 The Free Press, 1964.

15. Pèpe M. Tisserand-Perreer *Méthodes statistiques dans les sciences humaines,* Paris, Masson et Cie, 1962.

16. I. Marinescu, C. Moineagu, R. Niculescu şi V. Urseanu, *Elemente, de statistică matematică şi aplicaţiile ei,* Bucureşti, Edit. ştiinţifică, 1966.

17. Wright Lewall, *The method of path coefficients,* în: Anals of mathematical statistic, V, 1934.

18. John W. Tukey, *Causation, regression and path analysis,* în: Oskar Kempthorne ş. a., *Statistics and mathematics in biology,* Ames, Iowa State College Press, 1954.

19. E. J. Williams, *Regression analysis,* New York, Wiley, 1959.

20. A. G. Aganbegian, *Kolicestvenie metodi v sotiologii,* Moskva, Izd. „Nauka", 1966.

21. M. Şileţchi, L. Curta şi C. Ciochină, *Un model al migraţiei interzonale a populaţiei din R. S. România,* în: Studii şi cercetări de calcul economic şi cibernetică economică, nr. 3, 1969.

22. Herbert A. Simon, *Model of man,* New York, Wiley, 1957.

Scanned with OKEN Scanner

23. Vittorio Capecchi, *Causalità e correlazione nella problematica sociologica*, în : Studi di sociologia, 2, fasc. 3, 1964.

24. S. Novak, *Some problems of causal interpretation of statistical relationsship*, în : Philosophy of science, 27, 1960.

25. Gh. Mihoc şi N. Ionescu, *Bazele matematice ale programării lineare*, Bucureşti, Edit. ştiinţifică, 1965.

26. Gr. Osipov, *Metodika i technika statisticeskojobrabotki pervicnoj sociologiceskoj informacii*. Moskva, „Nauka", 1968 (Akad, Nauk SSSR, Institut filozofii).

27. Ştefan Szostkiewicz, *Proceduri i techniki badan socjologicznycb*, Warzawa, 1960, 4, Wojsk, Akademia Polit. in F. Dzierzynskiego, Katedra Socjologii.

Scanned with OKEN Scanner

5. MODELE DE ANALIZĂ FACTORIALĂ ŞI MULTIFACTORIALĂ

5.1. INTRODUCERE

Problema matematică fundamentală a analizei factoriale şi multifactoriale constă în reprezentarea unei serii de variabile, toate caracterizînd una şi aceeaşi populaţie, printr-un număr restrîns de factori independenţi. La rezolvarea acestei probleme o mare contribuţie au adus-o matematicienii şi statisticienii Charles Spearman, Louis Thurstone, Karl Holtzinger, Karl Pearson şi alţii [1].

Analiza factorială şi multifactorială şi-a găsit o primă aplicare în domeniul psihologiei, care, încă de la începutul secolului nostru, avea nevoie de un instrument statistic cu ajutorul căruia să se prelucreze un număr mare de teste psihologice. Un astfel de instrument l-a constituit analiza factorială şi multifactorială, cu ajutorul căreia o serie de teste psihologice au putut fi explicate printr-un număr redus de factori [2].

În ultimii ani, analiza factorială şi multifactorială a început să fie aplicată cu succes în cercetările sociologice, dovedindu-se un instrument valoros de analiză a fenomenelor şi proceselor sociale. De pildă, aceasta dă rezultate bune în studiul intensităţii acţiunii factorilor neobservabili, nemăsurabili sau care nu pot fi exprimaţi cantitativ.

Analiza factorială şi multifactorială este o ramură a analizei multidimensionale care se ocupă de studiul structurii interne a matricelor de corelaţie statistică.

Pentru analiza acestei structuri există două procedee : *a*) Procedeul de descompunere a matricei de corelaţie într-o mulţime de componente ortogonale egale ca număr cu

numărul variabilelor implicate în analiză. Acest procedeu permite evaluarea procentului din dispersia totală a variabilelor datorate fiecărei componente. Deşi se poate întîmpla ca un număr mic de componente să acopere o fracţiune mare a dispersiei, toate componentele sînt necesare în analiză. Procedeul de descompunere a matricei de corelaţie poate fi aplicat fără restricţii funcţiei de repartiţie a variabilelor analizate. *b*) Procedeul de analiză a variabilelor normate în scopul explicării matricei de covarianţă printr-un număr minim de factori. Aceşti factori rezultaţi din analiză constituie variabile ce nu pot fi identificate cu precizie, cunoscîndu-se numai numărul lor şi intensitatea cu care acţionează asupra variabilelor iniţiale.

În prezentul capitol nu ne propunem să descriem toate metodele de analiză factorială şi multifactorială, ci numai acelea care prezintă un interes deosebit pentru cercetările sociologice concrete.

5.2. METODA DE ANALIZĂ FACTORIALĂ A LUI SPEARMAN

Analiza factorială a fost concepută ca metodă de cercetare de Charles Spearman [3]. Această metodă se bazează pe materialul statistic rezultat din măsurători efectuate pentru n variabile care caracterizează o populaţie formată din N elemente. De exemplu, cele n caracteristici pot reprezenta principalii indicatori economici şi sociali care caracterizează N judeţe ale ţării sau principalele capitole din bugetele a N familii.

Calculînd coeficienţii de corelaţie între aceste caracteristici luate două cîte două, se constată că între ele există relaţii statistice mai mult sau mai puţin strînse.

Pentru evidenţierea factorilor care se găsesc la baza acestor relaţii, analiza factorială porneşte de la următorul model:

$$\dot{Z}_i = a_i F + e_i, i = 1, 2, \ldots, n.$$

Scanned with OKEN Scanner

unde :

Z_i — variabilele (caracteristicile) care fac obiectul analizei factoriale şi pe care le presupunem ca fiind normale, cu media 0 şi dispersia 1;

e_i — factorii specifici;

a_i — parametrul modelului (coeficient, constantă);

F — factorul comun (sistematic) sau variabilă ipotetică.

Cantităţile Z_i sînt observabile iar F, e_i şi a_i sînt inobservabile. Acestea din urmă însă pot fi deduse prin rezolvarea modelului matematic.

Problema analizei factoriale prezintă două aspecte [4] : *a*) verificarea unei ipoteze statistice referitoare la numărul factorilor ; *b*) estimarea parametrilor a_j.

Cînd între variabilele Z_1 şi Z_2 există o legătură liniară sub forma unui nor de puncte de formă alungită, în jurul unei drepte, avem cazul unui singur factor comun, care geometric este redus la o mulţime liniară, adică la punctele dreptei (fig. 43).

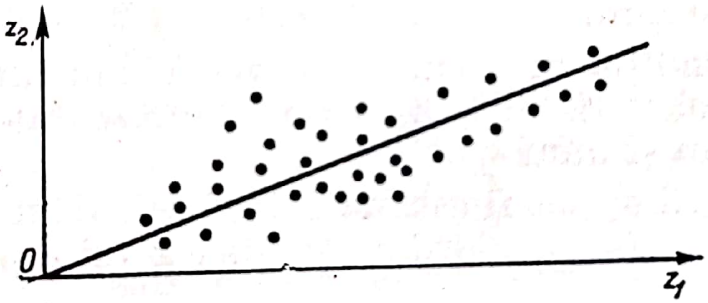


Fig. 43. — Reprezentarea geometrică a dependenţei liniare dintre doi factori (Z_1 şi Z_2)

Dacă avem trei variabile Z_1, Z_2 şi Z_3, pentru care valorile măsurate se reprezintă în spaţiul cu trei dimensiuni sub forma unui nor de puncte complet aplatizat (adică sub forma unui disc foarte plat) avem cazul unei reprezentări prin doi factori, care nu sînt altceva decît axele figurii plane care ajustează norul respectiv.

Generalizînd în continuare pentru un număr mare de variabile, norul de puncte indică aproximativ o sferă sau

o altă figură tridimensională, scoţînd în evidenţă dependenţa acestora de trei factori reprezentaţi geometric de axele sferii ş.a.m.d.

Există un criteriu datorat lui Spearman (1904) şi demonstrat în întregime de Garnet (1919), după care se poate aprecia că *matricea coeficienţilor de corelaţie* [5]:

$$R = \begin{pmatrix} 1 & r_{12} & r_{13} & r_{in} \\ r_{21} & 1 & r_{23} & r_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & r_{n3} & 1 \end{pmatrix}$$

admite un singur factor comun numai dacă variabilele tetrade ale acestei matrice se anulează.

O diferenţă tetradă a matricei R este o expresie de forma: $r_{ab}r_{cd} - r_{bc}r_{ad}$; $a, b, c, d, = 1, 2, \ldots, n$; $a \neq b \neq c \neq d$. În practică, această condiţie se verifică foarte rar; de aceea ea se consideră îndeplinită cînd aceste diferenţe sînt apropiate de zero.

Dacă condiţia menţionată se verifică în limite acceptabile, atunci variabilele Z_i pot fi reprezentate printr-un factor comun şi unul specific.

Parametrii a_i pot fi estimaţi dacă calculăm coeficienţii de corelaţie pe baza valorilor teoretice Z_{ij} şi Z_{hj}, după cum urmează [5]:

$$r_{ih} = \frac{1}{N}\sum_{j=1}^{N} Z_{ij}Z_{hj} = \frac{1}{N}\sum_{j=1}^{N}(a_iF_j + e_{ij})(a_hF_j + e_{hj}) =$$

$$= \frac{1}{N}\sum_{j=1}^{N}(a_i a_h F_j^2 + a_h e_{ij}F_j + a_i e_{hj}F_j + e_{ij}e_{hj}) =$$

$$= \frac{a_i a_h}{N}\sum_{j=1}^{N} F_j^2 + \frac{a_h}{N}\sum_{j=1}^{N} F_j e_{ij} + \frac{a_i}{N}\sum_{j=1}^{N} F_j e_{hj} + \frac{1}{N}\sum_{j=1}^{N} e_{ij}e_{hj}.$$

Din ipotezele de normalitate făcute asupra variabilelor F şi e, rezultă [5]:

$$\frac{1}{N}\sum_{j=1}^{N} F_j^2 = 1\,;\quad \frac{1}{N}\sum_{j=1}^{N} F_j\, e_{ij} = \frac{1}{N}\sum_{j=1}^{N} F_j\, e_{hj} = \frac{1}{N}\sum_{j=1}^{N} e_{ij}\, e_{kj} = 0.$$

Se obţine atunci:

$$r_{ih} = a_i \cdot a_h\,.$$

Parametrii a_i pot fi determinaţi cu ajutorul formulei [5]:

$$a_i = \left(\frac{r_{jh}\cdot r_{jl}}{r_{hl}}\right)^{\frac{1}{2}}$$

Cum Z_{ij} şi Z_{hj} sînt aleşi în mod arbitrar, se pot lua toate perechile posibile, iar a_i se estimează prin valoarea medie a celor C_{n-1}^2 estimaţii:

$$a_i^2 = \frac{1}{C_{n-1}^2}\sum_{\substack{h<l=1\\ hl\neq i}}^{n} \frac{r_{ih}\cdot r_{il}}{r_{hl}}$$

În cazul în care n este foarte mare, a_j se mai poate estima cu ajutorul expresiei [5]:

$$a_i^2 = \frac{\left(\sum_{h=1}^{n} r_{ih}\right)^2 - \sum_{h=1}^{n} r_{ih}^2}{2\left(\sum_{h<l=1}^{n} r_{hl} - \sum_{h=1}^{n} r_{ih}\right)},$$

unde:

Σr_{ih} este suma coeficienţilor de corelaţie dintre variabila x_i şi toate celelalte variabile;

Σr_{ih}^2 este suma pătratelor acestor coeficienţi;
Σr_{hl} este suma tuturor coeficienţilor de corelaţie.

Parametrul a_i reprezintă coeficientul de corelaţie între Z_i şi F, deoarece:

$$Z_i \,.\, F = a_i \,.\, F^2 + e_i \,.\, F,$$

de unde:

$E(Z_iF) = r_{Z_{i.F}} = a_i$, dacă acceptăm ipotezele de normalitate:
$E(F^2) = 1$; $E(e_i F) = r_{e_{i.F}} = 0$.

Factorul comun F este legat de celelalte variabile social-economice analizate. Determinarea valorii şi interpretarea acestui factor este extrem de dificilă şi nu întotdeauna posibilă.

Totuşi, dacă sîntem în măsură să efectuăm o analiză factorială, interpretarea factorului F trebuie făcută după criterii pur sociologice.

Analiza factorială ne permite să măsurăm intensitatea legăturii statistice dintre un factor observabil (Z_i) şi un factor neobservabil (F) sau chiar nemăsurabil. Această posibilitate pe care o oferă această metodă de analiză o face deosebit de utilă pentru cercetările sociologice [5].

5.3. UN EXEMPLU DE ANALIZĂ A FACTORULUI DE PROGRES TEHNIC ÎNREGISTRAT ÎN ECONOMIA NAŢIONALĂ

În studiul progresului tehnic înregistrat în economia naţională trebuie să adoptăm un sistem de indicatori sintetici ca de exemplu [6]:

a) Productivitatea muncii pe locuitor (z_1);
b) Fondurile fixe (z_2);
c) Volumul investiţiilor (z_3);
d) Venitul naţional pe cap de locuitor (z_4);
e) Energia electrică consumată în industrie (z_5).

Scanned with OKEN Scanner

Datele referitoare la dinamica acestor variabile le găsim în anuarul statistic al R.S.România pe anul 1968. Deoarece analiza factorială operează numai cu variabile standardizate (fără dimensiune), se pot utiliza atît date relative, cît şi date absolute, fără ca acestea să prezinte un inconvenient pentru aplicarea metodei [6].

Seriile din anuar vor fi transformate astfel:

$Z_{ij} = \frac{z_{ij} - \bar{z}_i}{\sigma_i}$; $i = 1, \ldots, 5$; $j = 1, \ldots, 10$, unde z_{ij} sînt datele brute pentru seria i în anul j:

$$\bar{Z}_i = \frac{\sum_{j=1}^{10} z_{ij}}{10}; \qquad \sigma_i = \sqrt{\frac{\sum_{j=1}^{10} (z_{ij} - \bar{z}_i)^2}{10}}.$$

Valorile astfel prelucrate ale variabilelor Z_i sînt înscrise în tabelul 18 [6]:

Tabelul 18

Anul	Z_1	Z_2	Z_3	Z_4	Z_5
1	−1,40	−1,32	−1,10	−1,09	−1,17
2	−1,16	−1,13	−0,93	−1,36	−1,00
3	−0,86	−1,00	−1,02	−0,86	−0,86
4	−0,44	−0,60	−0,88	−0,80	−0,66
5	−0,34	−0,30	−0,59	−0,34	−0,45
6	0,14	0,04	0,01	0,10	0,13
7	0,46	0,41	0,54	0,52	0,27
8	0,80	0,75	0,98	0,72	0,70
9	1,22	1,25	1,29	1,20	1,25
10	1,80	1,77	1,69	1,84	2,02

Coeficienţii de corelaţie dintre variabilele Z_i, calculaţi cu ajutorul formulei [6]:

$$r_{ij} = \frac{1}{10} \sum_{j=1}^{10} Z_{ik} Z_{jk},$$

Scanned with OKEN Scanner

sînt înscrişi în tabelul 19.

Tabelul 19

Z_j \ Z_i	Z_1	Z_2	Z_3	Z_4	Z_5
Z_1	1,000	0,989	0,973	0,983	0,984
Z_2	0,989	1,000	0,977	0,977	0,983
Z_3	0,973	0,977	1,000	0,984	0,991
Z_4	0,983	0,977	0,984	1,000	0,986
Z_5	0,984	0,983	0,991	0,986	1,000

Deoarece coeficienţii de corelaţie din tabelul 19 au valori apropiate de unu, aceştia reflectă legături foarte strînse între variabilele studiate.

Avînd în vedere coeficienţii de corelaţie din tabelul 19 ne punem problema dacă variabilele Z_i admit o reprezentare de forma [7]:

$$Z_i = a_i F + e_i; \quad i = 1, 2, \ldots, n.$$

În acest scop, vom verifica ipoteza nulităţii celor $3C_5^4 = 15$ diferenţe tetrade. Fie diferenţa $r_{12} r_{34} - r_{13} r_{24} = 0{,}989 \times 0{,}988 - 0{,}973 \times 0{,}977 = 0{,}0217$.

Calculînd pentru această diferenţă, eroarea probabilă de selecţie, cu ajutorul formulei de la 5.2, obţinem 0,0242. Eroarea probabilă de selecţie fiind mai mare decît valoarea diferenţei tetrade obţinute, putem neglija această eroare. Procedînd în mod analog pentru celelalte 14 diferenţe tetrade, vom observa că ipoteza lui Spearman este verificată statistic.

Folosind formula de calcul a parametrilor a_i^2 de la 5.2, obţinem: $a_1^2 = 0{,}982$; $a_2^2 = 0{,}979$; $a_3^2 = 0{,}979$; $a_4^2 = 0{,}982$; $a_5^2 = 0{,}991$.

Cunoscînd aceşti coeficienţi putem scrie sistemul de ecuaţii liniare:

$$Z_1 = 0{,}990\, F + e_1;$$

$$Z_2 = 0{,}985\, F + e_2;$$

$$Z_3 = 0{,}985\ F + e_3;$$
$$Z_4 = 0{,}986\ F + e_4;$$
$$Z_5 = 0{,}995\ F + e_5.$$

Putem trage concluzia că intensitatea legăturii statistice dintre factorii observabili (Z_i) și factorul comun nemăsurabil (F) este foarte puternică. Nu cunoaștem însă valorile pe care le poate lua factorul comun și nici interpretarea economică sau sociologică ce i se poate da. Pentru exemplul nostru se pare însă că factorul F poate fi interpretat ca progres tehnic [6].

5.4. UN EXEMPLU DE CONSTRUIRE A UNUI INDICATOR SINTETIC AL NIVELULUI DE TRAI AL NAȚIUNILOR

O aplicație deosebit de interesantă a analizei factoriale este construirea unui indicator sintetic al nivelului de trai al națiunilor, extrem de util în comparațiile internaționale [8]. Acest indicator se obține prin selectarea unor produse cheie, în economia unei națiuni care asigură o măsurare directă a bunăstării populației. Deoarece produsele selectate sînt măsurate mai mult în unități naturale decît în expresie valorică, transformarea conform cursurilor de schimb nu este necesară, iar problema dificilă a ponderii de schimb este evitată.

Pentru ca indicatorul sintetic al nivelului de trai să constituie o bază utilă în efectuarea comparațiilor pe plan internațional, el trebuie să prezinte o imagine complexă, care să reflecte cît mai fidel bunăstarea națională, definită prin condițiile reale de trai ale populației. Primele încercări în direcția alegerii unor variabile adecvate pentru caracterizarea nivelului de trai au fost făcute de M. K. Bennet [8]. Ulterior s-a construit un indicator sintetic al nivelului de trai de către Donald N. Niewiaroski [9].

Indicatorul sintetic al nivelului de trai s-a fundamentat pe o serie de variabile economice și sociale capabile să

Scanned with OKEN Scanner

caracterizeze cît mai complet conceptul de „nivel de trai". Aici indicatorul nivelului de trai apare ca o medie ponderată a acestor variabile [6].

Analiza factorială s-a utilizat pentru determinarea ponderilor tuturor variabilelor selecţionate. În acest scop s-au ales 14 indicatori economici şi sociali [9]:

a) *indicatori economici*

— proporţia forţei de muncă angajată în activităţi neagricole (x_1);
— numărul de telefoane pe un locuitor (x_2);
— consumul de energie electrică pe un locuitor (x_3);
— consumul de oţel pe un locuitor (x_4);
— producţia de ciment pe un locuitor (x_5);
— vehicule cu motor pe un locuitor (x_6);
— consumul de fibre pe un locuitor (x_7);
— proporţia de proteine în regimul alimentar (x_8);
— exportul pe un locuitor (x_9).

b) *indicatori sociali*

— numărul de medici şi dentişti pe un locuitor (x_{10});
— circulaţia ziarelor pe un locuitor (x_{11});
— numărul de aparate de radio pe un locuitor (x_{12});
— numărul de elevi ai şcolilor secundare (x_{13});
— cultură generală (x_{14}).

În cercetarea întreprinsă s-au strîns date din 87 de ţări.

Indicatorul nivelului de trai s-a exprimat cu ajutorul variabilelor menţionate astfel:

$$I = b_1 x_1 + b_2 x_2 + b_3 x_3 + \ldots + b_{14} x_{14},$$

unde:

I — indicatorul nivelului de trai;
b_i — ponderea fiecărei variabile;
x_i — variabila.

Metoda matematică de analiză factorială este utilizată pentru a determina ponderile b reducînd astfel cele

14 variabile la o singură componentă comună. Se asigură în acelaşi timp : *a*) maximizarea variantei indicatorului Y ; *b*) maximizarea sumei corelaţiilor pătrate dintre indicator şi variabilele selectate.

Maximizarea variantei indicatorului I indică o discriminare între ţările cu niveluri de trai diferite, iar maximizarea sumei corelaţiilor pătrate exprimă o proporţie justă a varianţei totale a tuturor variabilelor utilizate.

Transformînd datele de bază într-o formă standard, pentru o problemă cu trei variabile, ecuaţia de mai sus devine [8] :

$$i = w_1 z_1 + w_2 z_2 + w_3 z_3,$$

în care w_1, w_2 şi w_3 sînt ponderile pentru variabilele standardizate z. Aceste ponderi permit o variaţie însemnată a indicatorului, astfel încît să se diferenţieze ţările cu nivel de trai ridicat, mediu şi scăzut. În acest caz, varianţa indicatorului este :

$$\text{var } i = \frac{1}{n} \sigma_i^2 = g\,(w_1^2 + w_2^2 + w_3^2),$$

în care :

g este multiplicatorul lui Lagrange.

Suma corelaţiei pătrate este :

$$r_{z_1 i}^2 + r_{z_2 i}^2 + r_{z_3 i}^2 = g\,(w_1^2 + w_2^2 + w_3^2).$$

Întrucît suma ponderilor pătrate : $w_1^2 + w_2^2 + w_3^2 = 1$, prin maximizarea varianţei indicatorului se maximizează şi suma corelaţiilor pătrate.

În lucrarea lui D. N. Niewearoski s-a calculat *matricea de corelaţie* din tabelul 20.

Scanned with OKEN Scanner

Tabelul

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1,0	0,73	0,73	0,77	0,74	0,67	0,92
x_2	0,73	1,0	0,77	0,78	0,71	0,90	0,81
x_3	0,73	0,77	1,0	0,91	0,74	0,77	0,87
x_4	0,77	0,78	0,91	1,0	0,81	0,70	0,90
x_5	0,74	0,71	0,74	0,81	1,0	0,58	0,80
x_6	0,67	0,90	0,77	0,70	0,58	1,0	0,72
x_7	0,82	0,81	0,87	0,90	0,80	0,72	1,0
x_8	0,73	0,70	0,73	0,70	0,63	0,65	0,73
x_9	0,71	0,74	0,66	0,66	0,77	0,63	0,65
x_{10}	0,80	0,60	0,66	0,74	0,75	0,49	0,83
x_{11}	0,80	0,78	0,76	0,82	0,78	0,65	0,89
x_{12}	0;80	0,72	0,69	0,73	0,69	0,66	0,84
x_{13}	0,85	0,63	0,69	0,73	0,72	0,53	0,80
x_{14}	0,76	0,86	0,90	0,85	0,70	0,83	0,88

Ponderile obţinute prin analiza factorială se dau în tabelul 21 (pag. 228).

În tabelul 21 se vede că aproximativ 76 % din varianţa totală a variabilelor standardizate [9] se explică prin componenta principală. Ideal ar fi să se poată explica 100 %, ceea ce este posibil dacă s-ar lărgi grupul de variabile. De pildă, s-ar putea include în analiză şi unele variabile ca : situaţia locuinţelor, morbiditatea, durata medie a vieţii, mortalitatea infantilă, economiile populaţiei etc. care ar determina un procent mai mare din varianţă explicat cu ajutorul componentei principale.

Corelaţiile dintre indicator şi fiecare caracteristică sînt prezentate în tabelul 22 (pag. 228).

Valorile ridicate obţinute pentru toţi coeficienţii din tabelul 22 confirmă alegerea celor 14 variabile pentru măsurarea nivelului de trai şi atestă valabilitatea metodei folosite pentru determinarea indicatorului sintetic I. Prin urmare, în comparaţiile economice internaţionale poate fi aplicată analiza factorială la elaborarea sistemului de indicatori necesari.

Scanned with OKEN Scanner

20

x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}
0,73	0,71	0,80	0,80	0,83	0,53	0,76
0,70	0,74	0,60	0,78	0,72	0,63	0,86
0,73	0,66	0,66	0,76	0,69	0,69	0,90
0,70	0,66	0,74	0,82	0,73	0,73	0,85
0,63	0,77	0,75	0,78	0,69	0,72	0,70
0,65	0,63	0,49	0,65	0,66	0,53	0,83
0,73	0,65	0,83	0,89	0,84	0,80	0,88
1,0	0,67	0,69	0,72	0,59	0,77	0,74
0,67	1,0	0,50	0,73	0,64	0,61	0,60
0,69	0,50	1,0	0,75	0,68	0,84	0,71
0,72	0,73	0,73	1,0	0,84	0,80	0,77
0,59	0,64	0,68	0,84	1,0	0,79	0,74
0,77	0,61	0,84	0,80	0,79	1,0	0,72
0,74	0,60	0,71	0,77	0,74	0,72	1,0

5.5. UN EXEMPLU DE ANALIZĂ FACTORIALĂ A RELAȚIEI DINTRE CARACTERISTICILE SOCIAL-POLITICE ȘI PRODUSUL GLOBAL PE LOCUITOR

Se consideră că variabilele standardizate Z_{it} pot fi substituite cu ajutorul unei serii de variabile y_m ortogonale și necorelate ($m = 1, \ldots, n$), ale căror valori y_{mt} sînt legate de Z_{it} prin relații de forma [6]:

$$Z_{it} = \sum_{k=1}^{n} a_{ik} y_{kt} \quad (i, m = 1, \ldots, n).$$

În această ipoteză, coeficienții de corelație vor fi:

$$r_{ik} = \frac{\sum_{t=1}^{N} Z_{it} Z_{kt}}{N} = \sum_{s=1}^{N} \sum_{p=1}^{N} a_{is} a_{kp} \frac{\sum_{t=1}^{n} y_{st} y_{pt}}{N}$$

Tabelul 21

Nr. crt.	Variabila	Ponderile: combinată	economică	socială
1	Utilizarea forței de muncă neagricolă	0,0855	0,1254	
2	Telefoane %	0,0219	0,0341	
3	Consumul de energie %	0,0014	0,0017	
4	Consumul de oțel %	0,0150	0,0233	
5	Producția de ciment %	0,0123	0,0191	Neaplicabilă
6	Vehicule %	0,0246	0,0390	
7	Consumul de fibre %	0,5779	0,8488	
8	Calitatea regimului alimentar	0,1197	0,1802	
9	Exportul %	0,0157	0,0241	
10	Medici și dentiști %	2,4969	—	8,0845
11	Circulația ziarelor	0,0137	Neaplicabilă	0,0423
12	Înscrieri în școli secundare %	0,0780		0,2472
13	Știința de carte %	0,0540	—	0,0384
14	Aparate de radio %	0,0130	—	0,0384
	Varianța totală a variabilelor	140	9,0	5,0
	Varianța calculată a variabilelor	10,6049	6,9224	4,0580
	Procentul variației totale explicate prin conceptul de nivel de trai	76	77	81

Tabelul 22

Nr. crt.	Variabila	Coeficientul de corelație: combinată	economică	socială
1	Utilizarea forței de muncă neagricolă	0,90	0,87	
2	Telefoane %	0,88	0,91	
3	Consumul de energie %	0,90	0,91	
4	Consumul de oțel	0,91	0,92	
5	Producția de ciment %	0,86	0,86	
6	Vehicule %	0,80	0,84	
7	Consumul de fibre %	0,95	0,93	
8	Calitatea regimului alimentar	0,82	0,83	
9	Exportul %	0,78	0,82	
10	Medici și dentiști %	0,83	Neaplicabil	0,88
11	Circulația ziarelor %	0,91		0,92
12	Înscrieri în școli secundare %	0,86		0,90
13	Știința de carte	0,86		0,92
14	Aparate de radio %	0,91		0,87

Scanned with OKEN Scanner

Întrucît variabilele y_m sînt presupuse ortogonale şi necorelate, rezultă [6]:

$$\frac{\sum_{t=1}^{N} y_{st}^2}{N} = 1; \quad \frac{\sum_{t=1}^{N} y_{st}\, y_{jt}}{N} = 0.$$

Din aceste relaţii regăsim:

$$r_{im} = \sum_{s=1}^{n} a_{is} \cdot a_{ms}; \quad i, m = 1, \ldots, n.$$

În prealabil este necesar să stabilim dispersia variabilelor Z în funcţie de y_m [6]:

$$\sigma^2 = \frac{\sum_{t=1}^{N} Z_{it}^2}{N} = \sum_{m=1}^{n} \sum_{s=1}^{n} a_{im}\, a_{is} \sum_{t=1}^{N} \frac{y_m\, y_{st}}{N} \sum_{1}^{n} a_{im}^2 = 1.$$

Deoarece $\sum_{1}^{n} a_{im}^2 = 1$, dispersia totală a ansamblului de variabile Z_i este:

$$\sum_{1}^{n} a_{i1}^2 + \sum_{1}^{n} a_{i2}^2 + \ldots + \sum_{1}^{n} a_{in}^2 = n.$$

Contribuţia procentuală a componentei cu numărul de ordine m la dispersia totală a sistemului va fi:

$$\frac{\sum_{1}^{n} a_{im}^2}{n} 100 = \frac{g_m}{n} \cdot 100; \; m = 1, \ldots, n.$$

Scanned with OKEN Scanner

Pentru $m = 1$, coeficienţii a_{i1} se vor determina prin rezolvarea unei probleme de maxim [6]:

$$\sum_{i=1}^{n} a_{i1}^2 = \max.$$

$$\sum_{s=1}^{n} a_{is}\, a_{js} = r_{ij}.$$

Acest sistem este format din $p^2 + 1$ ecuaţii, dintre care p^2 sînt restricţii. De aceea, pentru rezolvarea acestui sistem de ecuaţii se foloseşte *metoda multiplicatorilor a lui Lagrange* [6]. Utilizînd această metodă, coeficienţii a_{i1} se vor determina din sistemul omogen:

$$\begin{aligned}
(1 - g_1)\, a_{11} + \quad & r_{12}\, a_{21} \quad + \ldots + \quad r_{1n}\, a_{n1} \quad = 0 \\
r_{21}\, a_{11} + (1 - g_1)\, a_{21} + \ldots + \quad & r_{2n}\, a_{n1} \quad = 0 \\
\ldots\ldots\ldots

ajutorul formulei :

$$a_{im} = \sqrt{\frac{g_m}{\sum_1^n \Delta_{im}^2}} \cdot \Delta_{im}; \quad i = 1, \ldots, n,$$

unde :

Δ_{im} este complementul algebric al elementului situat în linia i şi coloana m.

După ierarhizarea factorilor se pune problema determinării numărului de factori care trebuie păstraţi pentru analiza ulterioară. De obicei numărul factorilor este egal cu numărul rădăcinilor mai mari decît unitatea obţinută din ecuaţiile :

$$\begin{aligned}
&(1-g)\, a_{11} + r_{12}\, a_{12} + \ldots + r_{1p}\, a_{1p} = 0 \\
&r_{12}\, a_{12} + (1-g)\, a_{22} + \ldots + r_{2p}\, a_{2p} = 0 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&r_{1p}\, a_{1p} + r_{2p}\, a_{2p} + \ldots + (1-g) = 0
\end{aligned}$$

Spre exemplificare, vom examina succint legătura dintre caracteristicile social-politice şi produsul global pe cap de locuitor [9], apelînd la aplicaţiile analizei factoriale. În acest studiu, s-au considerat 22 de variabile independente (caracteristici social-politice) şi o variabilă dependentă (produsul global pe cap de locuitor), pentru un număr de 74 de ţări slab dezvoltate din Africa şi America de Sud.

Reproducem tabelul rezumativ al acestei analize (tabelul 23), în care sînt trecute variabilele şi ponderile celor patru factori comuni, precum şi procentul din variaţia fiecărei variabile datorată factorilor extraşi.

Scanned with OKEN Scanner

Tabelul 23

Nr. crt.	Explicaţia	Ponderea factorilor				g_i
		F_1	F_2	F_3	F_4	
1	Produsul global pe cap de locuitor	0,64	0,43	−0,22	0,12	0,661
2	Suprafaţa agricolă cultivată prin metode tradiţionale	0,83	−0,32	0,21	0,03	0,832
3	Caracterul organizaţiilor sociale de bază	0,86	0,26	−0,05	−0,04	0,819
4	Răspîndirea ştiinţei de carte	0,83	0,37	−0,08	0,09	0,840
5	Potenţialul mijloacelor de informare cu caracter de masă	0,85	0,34	−0,08	0,04	0,843
6	Gradul de omogenitate culturală şi etnică	0,72	−0,30	0,07	0,17	0,646
7	Importanţa clasei mijlocii indigene	0,63	0,31	−0,40	0,05	0,658
8	Modernizarea aspectului exterior al oraşelor	0,64	0,46	−0,33	0,13	0,753
9	Eficienţa instituţiilor democratice	0,36	0,79	−0,17	0,19	0,815
10	Libertatea opoziţiei politice şi a presei	0,27	0,86	0,40	0,08	0,829
11	Gradul de fracţionare a partidelor politice	0,33	0,78	−0,40	−0,30	0,810
12	Baza sistemului partidelor politice	0,42	0,68	−0,06	−0,11	0,657
13	Forţa mişcării muncitoreşti	0,29	0,71	−0,34	0,08	0,715
14	Forţa politică a armatei	0,38	−0,58	0,29	−0,36	0,690
15	Eficacitatea administraţiei	0,30	0,53	−0,48	−0,20	0,636
16	Gradul de centralizare a puterii politice	−0,03	−0,76	0,20	−0,07	0,617
17	Mobilitatea socială	0,42	0,14	−0,55	0,24	0,549
18	Puterea aristocraţiei indigene	0,03	0,15	0,82	−0,04	0,700
19	Extinderea naţionalismului şi a ideii unităţii naţionale	0,61	−0,05	−0,57	−0,01	0,694
20	Participarea conducerii în procesul dezvoltării	0,11	0,21	−0,75	0,29	0,696
21	Participarea statului la activitatea economică	0,27	−0,41	−0,48	−0,41	0,638
22	Tensiunea socială	−0,23	0,03	0,08	−0,84	0,771
23	Stabilitatea sistemului politic	0,04	0,08	−0,24	0,86	0,808

Scanned with OKEN Scanner

Pe baza rezultatelor cercetării efectuate s-a stabilit g_1 pentru produsul global pe locuitor :

$$g_1 = (0{,}64)^2 + (0{,}43)^2 + (-0{,}22)^2 (0{,}12)^2 = 0{,}661.$$

Rezultă că circa 66% din variaţia produsului global se datorează celor patru factori determinaţi.

Criteriile şi consideraţiile folosite în alegerea caracteristicilor nu sînt, în unele cazuri, semnificative. De pildă, omiterea unor caracteristici cum sînt, forma de proprietate asupra pămîntului, respectiv efectuarea sau neefectuarea reformei agrare şi limitele ei ; gradul de dependenţă faţă de capitalul străin etc. nu poate corespunde exigenţelor ştiinţifice, mai ales cînd în locul lor sînt introduse caracteristici nesemnificative. În plus, dacă unele caracteristici sînt măsurate în mod cu totul arbitrar (eficienţa administraţiei, forţa politică a armatei etc.), analiza factorială efectuată are lacune serioase [6].

5.6. UN EXEMPLU DE ANALIZĂ FACTORIALĂ A POZIŢIEI SOCIAL-ECONOMICE A AGRICULTORILOR

Analiza factorială s-a utilizat cu succes în studiul poziţiei social-economice a unui grup de agricultori din India [13]. Aplicînd metoda centroidei de analiză a factorilor cu rotaţie ortogonală, P a r e e k şi T r i v e d i au extras trei factori ai poziţiei social-economice[6] [14], [15] şi [16] :

— influenţa socială sau felul de viaţă ;
— casta sau ocupaţia ;
— pămîntul sau poziţia social-economică.

Eşantionul a fost alcătuit dintr-un grup de 312 familii dintr-un ansamblu de sate de lîngă Delhi. Informaţiile

[6] Punctele de vedere sau concluziile lui U. Pareek şi G. Trivedi nu trebuie să fie interpretate ca reprezentînd opinia autorilor.

Scanned with OKEN Scanner

culese au fost analizate cu ajutorul unei scale standardizate, constînd din nouă indicatori principali împărţiţi în indicatori secundari. Indicatorii principali au fost egal ponderaţi, iar pentru indicatorii secundari s-au utilizat ponderi diferenţiate.

Indicatorii principali folosiţi în scală au fost următorii : pămîntul (1), cultura (2), locuinţa (3), ocupaţia (4), casta (5), puterea gospodăriei (6), bunurile materiale (7), familia (8) şi participarea socială (9) [7].

Matricea de corelaţie a indicatorilor principali ai scalei se dă în tabelul 24.

Tabelul 24

Indicatorul	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
1		0,14	0,13	0,22	0,80	0,87	0,58	0,08	0,28
2	0,14		0,60	0,17	0,06	0,01	0,12	0,14	0,26
3	0,13	0,60		0,15	0,19	0,32	0,15	0,20	0,09
4	0,22	0,17	0,15		0,79	0,59	0,41	0,20	0,12
5	0,80	0,06	0,19	0,79		0,44	0,56	0,30	0,12
6	0,87	0,01	0,32	0,59	0,44		0,58	0,31	0,12
7	0,58	0,12	0,15	0,41	0,56	0,58		0,43	0,40
8	0,08	0,14	0,20	0,20	0,30	0,31	0,43		0,27
9	0,28	0,26	0,09	0,12	0,12	0,12	0,40	0,27	

Analiza factorială a fost efectuată pentru toţi cei nouă indicatori principali. Sarcinile factoriale ale acestor indicatori pe cele trei centroide, înainte de rotirea axelor, sînt date în tabelul 25.

În analiza factorială după procedeul indicat de Thurstone [4], Guilford [17] şi Thomson [18], comunalitatea fiecărui indicator a fost luată egală cu corelaţia cea mai puternică din coloană reflectîndu-se

[7] Sistemul de indicatori a fost elaborat pentru condiţiile social-economice specifice ale Indiei.

Tabelul 25

Indicatorul	Sarcinile factoriale			Comunalitatea
	I_0	II_0	III_0	h^2
0	1	2	3	4
1. Pămîntul	0,748	0,401	0,262	0,789
2. Cultura	0,394	—0,604	—0,251	0,583
3. Locuinţa	0,456	—0,481	—0,245	0,499
4. Ocupaţia	0,644	0,273	—0,462	0,703
5. Casta	0,763	0,424	—0,332	0,872
6. Puterea gospodăriei	0,772	0,358	0,245	0,784
7. Bunurile materiale	0,712	0,108	0,297	0,606
8. Familia	0,446	—0,199	0,184	0,272
9. Participarea socială	0,385	—0,219	0,280	0,275
Procentul de varianţă	37	14	9	

asupra unor indicatori reziduali factorilor pentru a maximaliza pe cei cu semnele pozitive.

Factorii au fost extraşi prin aplicarea *metodei „centroidei"* obţinîndu-se pentru cei trei factori sarcinile factoriale din tabelul 8 [4]. La determinarea numărului de factori extraşi a fost utilizat un criteriu empiric stabilit de T u c k e r, numit „Ø-ul lui Tucker" [14].
Regula este dată de relaţia :

$$Ø = \frac{\rho_{s+1}}{\rho_s} = \frac{n-1}{n+1},$$

în care n este numărul de indicatori din matricea de corelaţie, ρ_s este suma valorilor absolute ale reziduurilor după s factori şi cu diagonale corectate, iar ρ_{s+1} este suma valorilor absolute ale reziduurilor după $(s+1)$ factori cu valorile iniţiale sau cu diagonale corectate.

Dacă Ø este egal cu $\frac{n-1}{n+1}$ sau este mai mare decît această cantitate, s este în adevăr ultimul factor care

necesită extragerea. Pentru $n=9$, $\frac{n-1}{n+1}$ a fost egal cu 0,80. Valorile lui Ø după extragerea celor trei factori au fost 0,44, 0,66 şi respectiv 0,83. Cu toate că pentru cel de al treilea factor valoarea lui Ø a depăşit cantitatea $\frac{n-1}{n+1}$ (0,80), s-a hotărît să nu se extragă mai mult de trei factori.

Factorii centroidei au fost rotiţi pentru a se obţine semnificaţia lor psihologică şi matematică. În acest scop a fost aplicată rotirea ortogonală a axelor, luînd doi factori în acelaşi timp [18]. Poziţia noilor axe s-a stabilit pe baza maximalizării sarcinilor factorilor marcaţi prin 0 şi minimalizării sarcinilor negative (fig. 44). S-au efectuat două rotaţii în sensul acelor de ceasornic. Prima rotaţie a fost efectuată între I_0 şi II_0, unghiul de rotaţie fiind de 61° (fig. 44, înregistrarea 1). A doua rotaţie de 39°, a fost aceea a factorilor II_0 şi III_0 (fig. 44, înregistrarea 2).

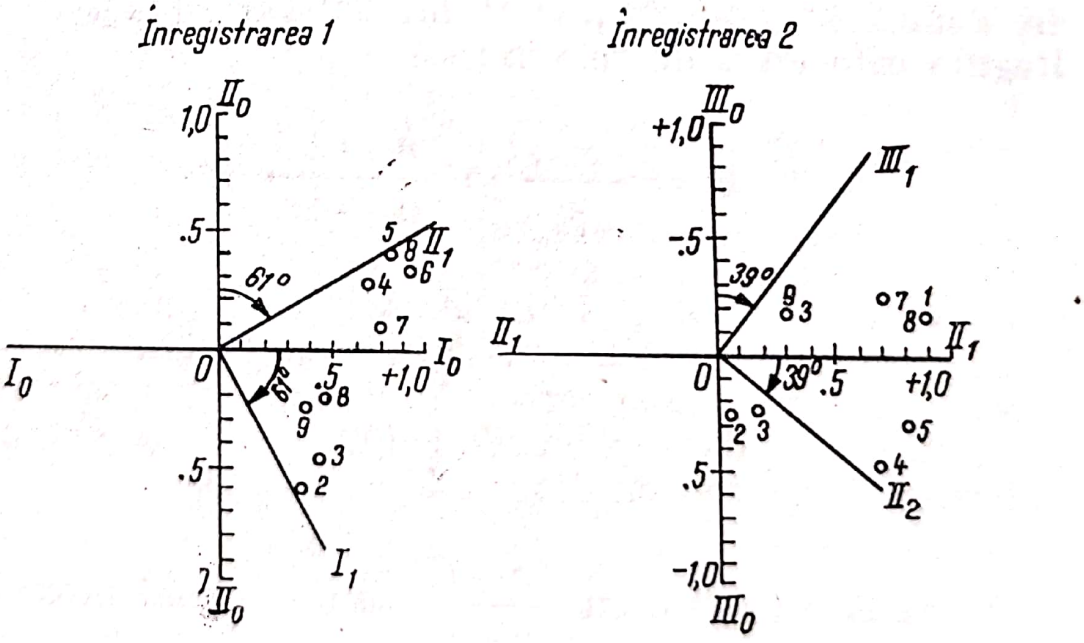


Fig. 44. — Rotirea ortogonală a axelor

Sarcinile factoriale au fost calculate cu ajutorul următoarelor formule :

$$L_1 \cos \theta - L_2 \sin \theta = I_1 ,$$

$$L_1 \sin \theta + L_2 \cos \theta = I_2 ,$$

în care L_1 şi L_2 sînt sarcinile iniţiale, înainte de rotirea factorilor respectivi, θ este unghiul cu care au fost rotite axele, iar I_1 şi I_2 sînt valorile sarcinilor noi pe prima şi respectiv pe cea de a doua axă.

Sarcinile finale ale indicatorilor pentru cei trei factori se dau în tabelul 26.

Tabelul 26

Indicatorul	Sarcinile factoriale			Comunalitatea (h^2)
	I_0	II_0	III_0	
0	1	2	3	4
1	0,012	0,459	0,738	0,790
2	0,720	0,198	−0,162	0,584
3	0,642	0,283	−0,086	0,499
4	0,073	0,831	0,079	0,702
5	−0,001	0,877	0,291	0,871
6	0,061	0,506	0,724	0,784
7	0,253	0,337	0,655	0,607
8	0,390	0,112	0,327	0,272
9	0,378	0,003	0,363	0,276
Procentul de variaţie	14	25	21	

Din ultima coloană a tabelelor 25 şi 26 rezultă că, în ambele cazuri, valorile h^2 sînt aproape aceleaşi.

Rotirea ortogonală a axelor a scos în evidenţă existenţa a trei factori : primul factor constă din indicatorii 2 (cultura), 3 (locuinţa), 8 (familia) şi 9 (participarea socială) ; al doilea factor constă din 4 (ocupaţia) şi 5 (casta) ; iar al

Scanned with OKEN Scanner

treilea factor constă din 1 (pămîntul), 6 (puterea gospodăriei) şi 7 (bunurile materiale). Aceşti trei factori care constituie poziţia social economică pot fi notaţi prin simbolurile A, B şi C.

Calculul variaţiei a arătat că factorul A explică 14% din varianţă, factorul B — 25% şi factorul C — 21% din varianţă. Acest studiu ajută numai la structurarea domeniului, nestabilind importanţa părţilor componente.

Factorul A prezintă două aspecte la fel de importante; *modul de viaţă* şi *influenţa socială*. Modul de viaţă este determinat de nivelul de cultură şi de locuinţă, iar influenţa socială — de cultură, locuinţă, familie şi de participarea socială. De aceea, se pare că este mai indicat să se identifice acest factor drept „*influenţă socială*".

În majoritatea cazurilor, persoanele cu cultură mai ridicată sînt foarte influente în sate, deoarece sătenii consultă pe cei care au o cultură ridicată şi le cer sfaturi. Cererea de sfaturi măreşte influenţa socială fiind de fapt o formă a acesteia.

O locuinţă este de asemenea o variabilă a influenţei sociale, deoarece o locuinţă mare şi bună nu indică numai poziţia socială înaltă a unei persoane din sat, ci ea oferă şi locul unde alţii pot veni să discute, să se distreze, să se întrunească etc.

O familie mare are un prestigiu social la sate, are un cerc mai larg de rude şi de prieteni. Prin urmare, o astfel de familie are contacte mai numeroase cu străinii, oferind condiţiile prielnice unei interacţiuni mai largi şi posibilitatea unei participări sociale sporite. Toate aceste caracteristici sînt surse de influenţă socială.

Factorul A are o oarecare sarcină la articolele de bunuri materiale. De pildă, radioul sau televizorul este o sursă de influenţă, deoarece lumea se adună în locuinţa posesorului pentru a audia şi respectiv viziona programele de radio şi televiziune.

Factorul B are sarcini mari la pămînt, ocupaţie, castă şi puterea gospodăriei. Dintre aceştia, *casta* şi *ocupaţia* au sarcinile cele mai mari. De aceea, factorul B poate fi

Scanned with OKEN Scanner

numit fie „castă", fie „ocupaţie". Acest factor are sarcini mici la indicatorii privind bunurile materiale şi locuinţa, mai mici la cultură şi familie, iar aproape nule la participarea socială.

Factorul C are sarcini mari la variabilele : pământul, puterea gospodăriei şi bunurile materiale, care indică poziţia economică. De aceea, acest factor poate fi numit „pământ", „poziţie economică" sau „venit". Dintre aceştia, pământul este articolul de bază, deoarece „puterea" gospodăriei nu are sens decît în legătură cu pământul.

Cei trei factori (cultura, venitul şi ocupaţia) au fost folosiţi ca elemente componente importante de către Kuppuswamy [20] pentru măsurarea poziţiei social-economice în zonele urbane. Un model similar de factori poate fi folosit şi pentru zonele rurale din apropierea unui oraş mare (ca de pildă Delhi). Acest model permite urmărirea creşterii influenţei urbane asupra satelor din vecinătatea oraşului. Studiul acestei influenţe necesită cercetări ample deoarece presupune că poziţia social-economică a unei regiuni este influenţată de gradul ei de urbanizare.

Rezultatele studiului poziţiei social-economice a locuitorilor din satele situate în vecinătatea oraşului Delhi au fost interpretate în sensul că poziţia social-economică nu este un concept unidimensional, ci multidimensional, cu trei dimensiuni sub forma celor trei factori analizaţi mai înainte. Conceptul de poziţie social-economică unidimensională este ambiguu, deoarece există o deosebire evidentă între factorii sociali şi cei economici. Deosebirea între aceste două categorii de factori nu este însă bine definită. De aceea, cei trei factori importanţi : influenţa socială sau nivelul de trai, casta sau ocupaţia şi pământul — poziţia economică sau venitul, caracterizează mai bine conceptul de poziţie social-economică.

5.7. ANALIZA STRUCTURII LATENTE

Sociologia ca ştiinţă cercetează realitatea socială, uneori şi dincolo de propriile sale manifestări.

Scanned with OKEN Scanner

De obicei se consideră că o realitate socială este *latentă* atîta timp cît ea nu poate fi cunoscută numai din datele manifeste [21].

În cercetările sociologice se face o distincţie evidentă între conceptele de *manifest* şi *latent*; de pildă, dacă prin *conceptul de manifest* vom înţelege toate comportamentele care sînt direct observabile, prin *conceptul de latent* vom înţelege pe cele inobservabile, adică cele explicative.

Cercetarea concomitentă a unor fenomene manifeste şi a unor realităţi latente prezintă interes numai în cazul în care cunoaşterea ştiinţifică se ameliorează într-o asemenea măsură încît permite efectuarea unor previziuni. Desigur că efectuarea unor previziuni fundamentate ştiinţific n-ar fi posibilă fără introducerea unei unităţi de măsură adecvate. Prin introducerea unei unităţi de măsură, conceptului de latent i se poate substitui noţiunea precisă de „continuum latent".

În general, culegerea datelor manifeste se face printr-un test de cîteva întrebări, conceput astfel încît să permită tratarea unui ansamblu de observaţii calitative sistematic unite. Fiecare întrebare pusă tuturor subiecţilor populaţiei alese este astfel concepută încît subiectul să poată răspunde prin „da" sau „nu". Acest fel de formulare al întrebărilor este propriu tipului de întrebări (itemuri) dihotomice. În alte situaţii, se pot folosi şi cuplurile contrare: „adevărat-fals", „favorabil-defavorabil", „de acord — în dezacord" etc.

Studiul distribuţiei de frecvenţă a răspunsurilor la diverse itemuri constituie scopul cercetării, iar testul astfel alcătuit încît densitatea de probabilitate a unui răspuns afirmativ la toate itemurile ansamblului să fie egală cu produsul densităţilor probabilităţilor de răspuns afirmativ la fiecare dintre itemuri luate separat — constituie mijlocul de a ajunge la acest scop. Desigur că într-o analiză aprofundată, scopul şi mijloacele trebuie unite.

Cercetarea distribuţiilor de frecvenţă şi a curbelor de probabilitate nu ne spune nimic însă despre *clasele latente*, îndeosebi despre efectul fiecăreia dintre ele. Dar putem

Scanned with OKEN Scanner

considera caracteristicile cercetate ca *necunoscute*, folosite într-un anumit număr de *ecuații ale structurii latente*. Desigur că membrii din stînga ai ecuațiilor vor fi formați din caracteristicile cunoscute, adică din datele manifeste, cum sînt probabilitățile de răspuns la un item : $p_1, p_2, p_3, \ldots, p_n; p_{12}; p_{13}; p_{1m}$ etc. Este evident că numărul de ecuații va fi funcție de numărul itemurilor utilizate în test. Dacă m este acest ultim număr se poate arăta că numărul parametrilor manifești va fi egal cu 2^m, cifrele 1, 2, 3, ..., m fiind indicii itemurilor. De exemplu, cu 3 itemuri vom avea deci : $2^3 = 8$ ecuații ; cu 4 itemuri : $2^3 = 16$ ecuații etc.

Membrii din stînga ecuațiilor vor fi obținuți prin numărarea răspunsurilor afirmative la itemul 1, 2, 3, ..., n, apoi a răspunsurilor afirmative în același timp la 1 și la 2 etc., după care se împart la efectul total, pentru a obține valorile $p_1, p_2 \ldots p_{12}$ etc.

Membrii din dreapta ecuațiilor vor fi alcătuiți în primul rînd din parametrii manifești p_i ai primului nivel de stratificare. Fiecare parametru p_i constituie proporția celor care răspund afirmativ la itemul i din ansamblul populației testate și, în consecință, din întreaga porțiune a continuumului pe care se distribuie. Deoarece frecvența se exprimă prin probabilitate, putem scrie :

$$p_i = \int_{-\infty}^{+\infty} f_i(x) \cdot \varphi(x) \partial x,$$

unde : $f_i(x)$ — funcția de frecvență corespunzătoare curbei de probabilitate,

$\varphi(x)$ — frecvența populației cu același grad de „moral" (de exemplu : moralul muncitorului).

Vom repartiza populația testată de-a lungul unui „continuum". Existența acestui „continuum" se poate stabili numai dacă construim distribuția frecvențelor, ordonînd populația de indivizi care au același grad de „moral" raportată la populația totală a testului, pe axa

Scanned with OKEN Scanner

absciselor x. Apriori, ne putem imagina că toate valorile lui x sînt atribuite unor indivizi ai populaţiei, cel puţin între o valoare minimă şi o valoare maximă. Între aceste limite se repartizează sau se distribuie populaţia testată de-a lungul acestui „continuum", ca în figura 45.

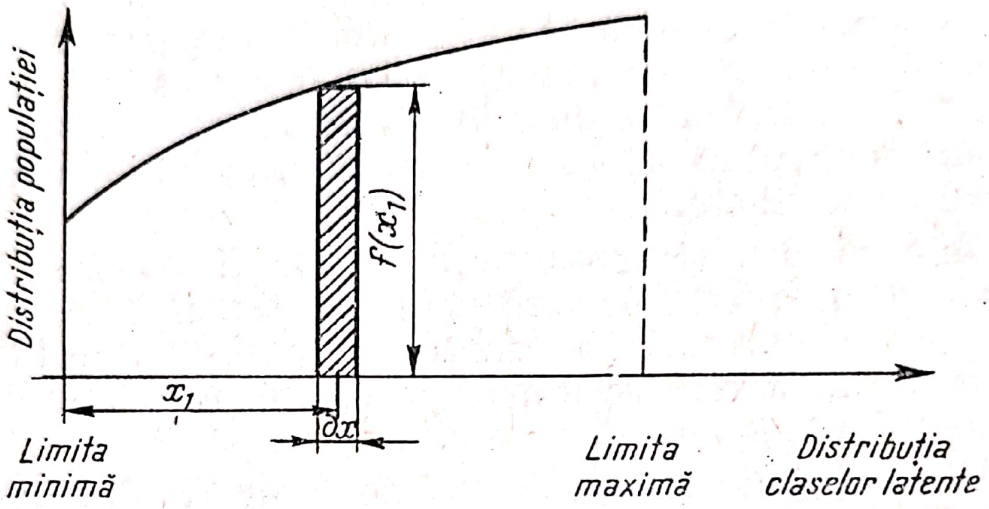


Fig. 45. — Repartizarea populaţiei testate de-a lungul unui „continuum"

Suprafaţa haşurată din fig. 45 se consideră de lăţime neglijabilă (∂x) şi de înălţime $\varphi(x_1)$, adică de $\varphi(x_1)\,\partial x$. Populaţia totală se distribuie pe întreg „continuumul" a cărui suprafaţă va fi: $\int \varphi(x)\partial x$.

Dacă trecem la un nivel de stratificare superior, unde frecvenţele manifeste sînt afectate de mai mulţi indici, putem scrie: $p_{1,\,2,\,3,\,\ldots,\,m} = p_1, p_2, p_3, \ldots, p_m$. Aceste probabilităţi pot fi scrise şi astfel: $p_{i,j,k} \ldots = \int_{-\infty}^{+\infty} f_i(x) \cdot f_j(x) \cdot f_k(x) \ldots \varphi(x)\partial x$.

Ecuaţiile prezentate sub această formă generală nu sînt utilizabile. Pentru a le face utilizabile este necesar să formulăm noi ipoteze. De pildă, să presupunem mai întîi, că densităţile de probabilitate sînt polinoame de forma:

$$f_i(x) = a_i + b_i x + c_i x^2 + \ldots + z_i x^n.$$

Dacă considerăm forma liniară a ecuațiilor $f_i(x)$, pentru $j=1$, vom avea :

$$f_i(x) = a_i + b_i x.$$

În acest caz, distribuțiile de probabilități vor avea forma :

$$p_i = \int_{-\infty}^{+\infty} (a_i + b_i x)\ \varphi(x)\ \partial x.$$

Scoțînd constantele în afara integralei, obținem :

$$p_i = a_i \int_{-\infty}^{+\infty} \varphi(x)\ \partial x + b_i \int_{-\infty}^{+\infty} x\ \varphi(x) \partial\ x.$$

Or, $\int \varphi\ (x)\ \partial x$ reprezintă populația totală, adică suma probabilităților (frecvențelor) care este egală cu 1. Pe de altă parte, $\int x \cdot \varphi(x)\ \partial x$ este definit ca moment de ordinul 1, pe care-l vom nota cu $M(1)$. Ecuația precedentă se va scrie acum :

$$p_i = a_i + b_i \cdot M_{(1)}.$$

În mod analog, o ecuație de grad superior se va scrie :

$$p_i = a_i + b_i \cdot M_{(1)} + c_i \cdot M_{(2)} + \ldots + z_i \cdot M_{(n)}.$$

La cel de-al doilea nivel de stratificare, dacă curbele iternurilor sînt de gradul I, vom avea :

$$p_{ij} = \int_{-\infty}^{+\infty} (a_i + b_i x)\ (a_i x + b_j x)\ \varphi(x) \partial x.$$

Scanned with OKEN Scanner

După o serie de transformări similare cu cele efectuate mai înainte, obţinem :

$$p_{ij} = a_{ij} + (a_i b_j + a_j b_i)\ M_{(1)} + b_i b_j\ M_{(2)}\,.$$

În cazul general vom avea :

$$p_{ij} = a_i \cdot a_j + (a_i b_j + a_j b_i)\ M_{(1)} + (a_i c_j + b_i b_j + c_i a_j)\ M_{(2)} + \\ + \ldots + z_i\, z_j M_{(2n)}.$$

Ecuaţiile structurii latente scrise mai sus permit analiza datelor manifeste. Membrii din dreapta ai ecuaţiilor constituie *modelul matematic*. Acest model va varia, evident, după structura socială sau psihosocială studiată, pe care trebuie s-o reflecte cu fidelitate.

În ecuaţiile structurii latente, momentele şi coeficienţii figurează ca necunoscute şi pot fi determinaţi prin introducerea ipotezelor simplificatoare propuse de Lazarsfeld [8].

5.8. CONCLUZII

Exemplele prezentate în acest capitol sînt departe de a epuiza domeniile de aplicare ale analizei factoriale. Din literatura de specialitate recentă rezultă că nu numai aplicaţiile ei, dar şi teoria acestei importante metode statistico-matematice este în continuă dezvoltare şi perfecţionare.

În prezent, se pun la punct noi metode mai complexe de analiză factorială, care înlătură unele restricţii impuse variabilelor (de pildă, liniaritatea, normalitatea acestora etc.), lărgind pe această cale posibilităţile de aplicare.

---

[8] Problemele de analiză a structurii latente sînt tratate în [22] şi [23].

Scanned with OKEN Scanner

Analiza factorială este, de fapt, cel mai adecvat instrument pentru studiul influenţei pe care o exercită asupra proceselor sociale factorii de natură diferită : economici, geografici, psihologici, biologici, culturali, istorici etc. De aceea, utilizarea analizei factoriale şi în cercetarea sociologică cere o strînsă colaborare între economişti, sociologi, statisticieni, psihologi şi matematicieni.

Numai prin realizarea acestei colaborări se poate analiza corect problema, alegînd în mod judicios variabilele şi interpretînd just rezultatele obţinute prin metodele cele mai adecuate de estimare a parametrilor şi de verificare a ipotezelor.

Analiza factorială a variabilelor economice şi sociale, cu toată precizia aparatului matematic utilizat, poate duce la concluzii greşite dacă nu se bazează pe o concepţie teoretică ştiinţifică, dacă nu este îmbinată cu o analiză complexă a variabilelor economice, sociale, geografice, biologice, psihologice etc. [24].

BIBLIOGRAFIE

1. T. W. Anderson, *An introduction to multivariate statistical analysis*, New York, Wiley, 1958.
2. G. H. Thomson, *L'analyse factorielle des aptitudes humaines*, trad. Naville, Paris, Presses Universitaires, 1950.
3. Charles Spearman, *General inteligence, objectively determined and measured*, în : American Journal of Psychology, 15, 1904.
4. Louis L. Thurston, *Multiple factor analysis*, Chicago, Chicago University Press, 1947.
5. Harman şi Harrary, *Modern factor analysis*, Chicago, Chicago University Press, 1960.
6. C. Ionescu, *Metode statistice de cercetare a corelaţiilor în economie*, Bucureşti, Edit. Academiei, „Bibliotheca Oeconomica" 3, 1967.
7. James S. Coleman, *Introduction to mathematical sociology*, Glencoe, Illinois (S. U. A.), The Free Press, 1964.
8. M. K. Bennet, *International Disposition in Consumation Levels*, în : American Economic Review, sep. 1951.

Scanned with OKEN Scanner

9. H. Donald şi W. Niewearoski, *Nivelul de trai al naţiunilor* (traducere). În : Estatistica, Journal of the Inter-American Statistical Institute, Whashington Dc. March, 23, nr. 86, 1965, p. 3—31.
10. Udai Pareck şi G. Trivedi, *Reability and Validity of Rural Socieconomics Status*. În : Scole & Indian Journal of Applied Psychology, 1964.
11. W. A. Gibson, *Three multivariate models: factor analysis latent structure analysis and latent profile analysis*, în : Psychometrika, 24, 1959.
12. N. Rancu şi L. Tövissi, *Statistica matematică cu aplicaţii în producţie*, Edit. Academiei, 1963.
13. Udai Pareck şi G. Trivedi, *Factor Analysis Socioeconomic status of Farmers in India*, în : Rural Sociology, 30, nr. 5, 1965.
14. G. Trivedi şi Udai Pareck, *Socioeconomic Status Scole (Rural), Form and Manual*, Delhi, Manasayan, 1964.
15. Udai Pareck şi G. Trivedi, *Norm for Delhi Area of Socioeconomic Status Scole (Rural)*, în : Buniayadi Talin, 6, 1963.
16. G. Trivedi şi Udai Pareck, *Categorisation of Rural Socioeconomic Groups*, în : Indian Journal of Social Work, 24, 1964.
17. J. P. Guilford, *Psychometric Methods*, New York — McGraw-Hill Book Co., 1954.
18. G. H. Thomson, *The Factorial Analysis of Human Ability*, London, University of London Press, 1955.
19. B. Tucker, *Introduction to Factor Analysis*, New York, D. Van Nostrand Co., 1954.
20. B. Kuppuswamy, *Socioeconomic Status Scole (Urban)*, Delhi, Manasayan, 1962.
21. Paul F. Lazarfeld, *Latent structure analysis*, în : Koch Sigmund, Psychology, *A study of a science*, vol. 3, New York, McGraw-Hill, 1959.
22. *Studies in Social Psychology*, IV, în : World War, 2.
23. *Mathematical Thinking* in the *Social Sciences*.
24. А. Н. Ефимова, *Проблемы оптимального планирования*, Москва, Изд. «Экономика», 1966, p. 146—190.

Scanned with OKEN Scanner

ÎNCHEIERE

În ultimii douăzeci de ani, noţiunea de sistem a căpătat o utilizare din ce în ce mai largă, extinzîndu-se mai întîi în domeniul tehnic, apoi în cel economic, pentru ca recent să-şi croiască drum în domeniul social. Urmărindu-se o tratare globală, s-a ajuns la situaţia de a se considera societatea ca sistem, punîndu-se adeseori în evidenţă subsistemele ei şi relaţiile dintre subsisteme.

Într-o accepţiune generală, prin *sistem* se înţelege o totalitate de elemente ordonate, legate între ele prin relaţii. Reprezentarea formală a acestor relaţii se numeşte *structura sistemului*. În acest caz, prin *dinamica sistemului* se înţelege ansamblul elementelor dinamice şi a relaţiilor existente în structuri dinamice, unde noţiunea de *dinamică* implică dependenţa de timp.

În realitate, dinamica sistemelor nu este doar rezultanta unor procese izolate, care se desfăşoară pe traiectoriile unor curbe individuale. Sistemele concrete antrenează elemente geografice, biologice, demografice, economice, politice, psihologice etc. Datorită interferenţei acestor elemente de natură diferită, starea viitoare a acestor sisteme nu poate fi studiată în mod fragmentar, izolat, ci numai în ansamblu. Prospectiva socială, „asamblînd" rezultatele obţinute în sectoare separate se constituie în fapt ca o ştiinţă de sinteză. De aceea, această ştiinţă apelează frecvent la descrierea sistemelor şi la urmărirea dinamicii lor, pentru ca, pornind de la sistemul social, să se obţină variante ale stării globale a sistemului, privite în mod unitar. Este evident că, într-o asemenea viziune, simpla corelare a tendinţelor izolate, pe compo-

Scanned with OKEN Scanner

nente, este insuficientă. Această tratare analitică nu numai că nu elimină analiza factorilor determinanţi, a dinamicii lor în timp, ci, mai mult chiar, devine inevitabilă.

Prospectiva socială ca ştiinţă, tinzînd să „reunifice" sistemul social din părţile sale componente, apelează consecvent la *metoda modelării multiple*. Această metodă se bazează pe studiile analitice separate ale unor procese sociale, vizînd urmărirea dinamicii fiecărui element al sistemului social, deci a întregii structuri sociale, asupra căreia se obţin mai multe imagini globale posibile. Dacă se vor introduce în studiul procesului social şi elementele care intervin imediat în dinamica sa, pot fi obţinute reprezentări ce vor acoperi realitatea sistemului într-un mod specific, cu numeroase interferenţe.

Metoda de previziune a sistemelor sociale care se fundamentează pe teoria generală a sistemelor, elaborată acum patru decenii, încearcă să considere realitatea socială ca un tot unitar. Această încercare însă se află în contradicţie cu modul de însumare mecanică a elementelor componente ale acestor sisteme.

Aşadar, sistemul social este descris prin elemente legate într-o structură socială, elemente care nu sînt disjuncte. De aceea, analiza fiecărui element implică de fapt considerarea întregului sistem în raport cu elementul fixat. Evident, un asemenea mod de abordare revine, logic, la construirea de subsisteme ale sistemului social luat în studiu, putînd să se definească prin mijlocire elementul urmărit. Prin urmare, analiza evoluţiei unui element implică o modelare multiplă a realităţii sociale, adică din puncte de vedere multiple, asociate diferitelor elemente luate în studiu. Fără îndoială că o asemenea modelare creează o viziune globală a sistemului social, contribuind la înţelegerea complexă şi multiplă a ideii de model dinamic al realităţii sociale.

În prezent, ritmul dezvoltării sociale a devenit atît de rapid încît previziunile sociale pe termen lung constituie o cerinţă imperioasă, dacă vrem să facem faţă cu succes condiţiilor de viaţă din viitor. Pentru aceasta, previziunile

Scanned with OKEN Scanner

sociale elaborate vor trebui să fie centrate pe mai multe idei, pe o imaginaţie creatoare, avîndu-se în vedere faptul că sistemele sociale viitoare vor fi constituite dintr-un complex de elemente economice, politice, culturale, psihologice etc., aflate într-o schimbare permanentă.

De multe ori, analiza unui sistem social conduce la construcţia unor modele multiple ale realităţii sociale, capabile să ofere o viziune globală a sistemului luat în studiu, devenind numai astfel operantă. Aceste modele sînt simbolice, adică logice sau matematice, de evoluţie a sistemelor sociale, avînd un rol explicativ-predictiv prin excelenţă.

Un *model explicativ-predictiv* al dinamicii sistemului social luat în studiu este o reprezentare, adesea simplificată, aproximativă, a unui domeniu al realităţii sociale, care păstrează evident, dintr-un punct de vedere stabilit, structura şi procesul social asociat domeniului considerat, privite în unitatea lor dialectică.

În cuprinsul lucrării am avut în vedere faptul că complexitatea dinamicii sistemelor sociale implică o abordare unitară a elementelor economice, politice, culturale, demografice, geografice, psihologice etc., ca un întreg, dacă nu vrem să omitem factori importanţi, a căror acţiune ar putea să „răstoarne" previziunile noastre.

Într-un mod mai explicit, unitatea dialectică dintre structură şi proces constituie un sistem dinamic ale cărui elemente se comportă ca un întreg. Într-un astfel de sistem, o structură ne conferă aspectele statice, pe cînd procesul ne furnizează aspectele dinamice.

O *structură* socială poate fi privită ca un ansamblu unitar de relaţii dintre elemente de natură diferită (economice, culturale, demografice, geografice, politice, psihologice etc.). Fără îndoială că aceste relaţii determină menţinerea fiecărui sistem ca o unitate, ca un întreg. De pildă, există o mulţime de relaţii de dependenţă între elemente economice, culturale, demografice, geografice, politice etc., care pot fi reunite într-o structură concretă.

Un *proces* social poate fi înţeles ca o transformare, schimbare sau evoluţie permanentă în timp a elementelor

structurii sociale în funcţie de care poate fi analizat trecutul şi prefigurat viitorul. Aceste transformări permanente se datorează unui ansamblu de factori previzibili sau imprevizibili. De exemplu, procesul de urbanizare implică o transformare şi perfecţionare permanentă a acţiunilor umane, sub impulsul factorilor de natură diferită : economici, social-culturali, psihologici etc.

Diferenţa dintre două structuri, aparţinînd unor perioade diferite de timp, poate fi privită ca o *mobilitate structurală*.

Spre deosebire de *analiza statică* a structurilor, care surprinde doar unele momente ale mobilităţii structurale, *analiza dinamică* ne semnalează evoluţia acestor structuri de-a lungul istoriei.

Avînd în vedere faptul că mobilitatea structurală este un fenomen esenţialmente structural, în sensul că mobilitatea unui element al structurii depinde de mobilitatea celorlalte elemente din sistem, putem compara mai multe structuri complete, destul de apropiate în timp, apte să evidenţieze tendinţele, să dezvăluie schimbările mai îndepărtate.

O analiză adîncită a mobilităţii structurale implică :

a) studiul transformării şi schimbării diferitelor elemente, în măsura în care ele aparţin structurii luate în studiu ;

b) cercetarea relaţiilor de dependenţă existente între diferitele elemente ale structurii, transformarea şi schimbarea lor ;

c) examinarea tendinţei trecute, prezente şi viitoare a transformării şi schimbării structurii sociale luată în totalitate.

Explorarea viitorului trebuie să aibă în vedere, înainte de toate, o analiză a factorilor determinanţi ai evoluţiei sistemelor, în funcţie de care să se contureze dezvoltarea lor pe termen lung.

Trecutul şi prezentul sînt necesare pentru cunoaşterea viitorului, dar ele nu sînt suficiente pentru „stăpînirea"

Scanned with OKEN Scanner

lui. Pentru a fi stăpînit un eveniment trebuie să cunoaştem factorii săi determinanţi, asupra cărora se poate acţiona în mod conştient.

Viitorul considerat pe termen lung nu poate fi privit ca o simplă prelungire a trecutului şi prezentului, datorită transformărilor permanente, a schimbărilor continue care intervin în condiţiile de existenţă, de cunoaştere şi de acţiune ale oamenilor, trăsătură fundamentală a vieţii lor.

Structurile sociale, începînd cu cele economice, trecînd prin cele demografice şi profesionale, pînă la cultură, ştiinţă etc., cunosc o evoluţie permanentă. Este bine ca de factorii determinanţi ai acestei evoluţii să ne ocupăm mai mult. Numai în felul acesta vom izbuti să elaborăm previziuni asupra schimbărilor posibile în viitor, pe baza cunoaşterii legilor obiective cărora le este subordonată mobilitatea structurală, al cărei ritm de evoluţie nu trebuie să găsească oamenii şi comunităţile umane nepregătite de a domina complexitatea, complicarea şi accelerarea dezvoltării sociale.

În acest scop, vor trebui create *modele matematice* care să simuleze perioadă cu perioadă dinamica socială sau anumite aspecte ale acestei dinamici : modele de creştere urbanistică, de simulare a consumului populaţiei, de localizare a industriilor şi a zonei destinată pentru locuinţe etc.

Modelarea matematică aplicată în cercetările sociologice prezintă o serie de avantaje evidente :

a) oferă o reprezentare intuitivă şi totuşi riguroasă a procesului social analizat ;

b) asigură verificarea, prin analogie, a consistenţei logice a teoriei şi ipotezei ;

c) contribuie la verificarea corectitudinii procesului de gîndire ;

d) favorizează elaborarea prognozelor sociale şi fundamentarea deciziilor optime ;

e) facilitează descoperirea unor relaţii şi legităţi la care nu am fi ajuns uşor pe alte căi.

Scanned with OKEN Scanner

Metoda modelării matematice, despre semnificaţia căreia vorbesc o serie de cercetări ştiinţifice remarcabile, nu constituie o descoperire nouă în ştiinţă. Doar denumirea de „modelare" pentru un anumit fel de reprezentare a unor procese sociale este de dată relativ recentă.

Preocupările oamenilor de ştiinţă de a utiliza metoda modelării matematice este strîns împletită cu întreaga istorie a dezvoltării ştiinţei, noutatea constînd doar în aplicarea sa riguroasă, dovedindu-se fertilă în dezvoltarea cunoaşterii progresului social.

Într-o accepţiune mai largă, modelul matematic, în măsura în care descrie în limbaj matematic un proces social oarecare, este construit totdeauna prin intermediul unei teorii. În acelaşi timp însă, un astfel de model serveşte şi ca verigă între teorie şi obiectul cercetat.

Relaţia *teorie — model — obiectul cercetat* apare sub două aspecte : unul *logic* şi altul *gnoseologic*. Din punct de vedere logic se cere realizarea unei corespondenţe între teorie şi model, gnoseologic se cere asigurarea corespondenţei dintre model şi obiectul cercetat.

În general, dacă teoria corespunde cu modelul iar acesta din urmă cu obiectul de cercetat, corespondenţa dintre teorie şi obiectul cercetat este asigurată. Această corespondenţă, însă, va trebui verificată în practică. Modelul pentru a avea utilitate în cercetările ştiinţifice trebuie să fie izomorf cu obiectul său, adică să aibe aceeaşi structură. De fapt, verificarea existenţei unui izomorfism între model şi obiectul la care se referă, echivalează cu verificarea valabilităţii obiective a modelului.

Confruntarea modelului matematic cu obiectul cercetat (procesul social) este practic imposibilă pe cale experimentală, deoarece nu putem reproduce procese sociale în condiţii de laborator. Aceasta nu înseamnă totuşi o absolută imposibilitate a experimentării, dacă avem în vedere faptul că unele fenomene sau procese sociale din trecut pot fi teoretic „modelate matematic" şi confruntate cu situaţii de dată mai recentă.

Scanned with OKEN Scanner

Noile descoperiri se pot face pe baza unor simulări pe modele matematice, fără să acţionăm nemijlocit asupra obiectului supus cercetării, de la care dorim să obţinem informaţii noi. Avantajul experimentării pe modele matematice este evident în cazul în care cu obiectul cercetării nu pot fi efectuate experienţe directe. Fără îndoială că procesul social luat în studiu depinde de mai multe variabile, între care există relaţii matematice, apte să permită efectuarea tuturor transformărilor posibile pe model, ale căror rezultate, dacă sînt considerate ipoteze şi dacă sînt verificate ulterior, pot releva fapte nebănuite despre obiectul cercetat.

Efectuînd asemenea experimente pe modele, putem face *previziuni*, intervenind conştient pentru modificarea favorabilă a procesului social studiat, ca de pildă: diminuarea fluxurilor migratorii ale populaţiei printr-o dezvoltare social-economică raţională a diviziunilor teritorial administrative ale ţării.

Desigur că transformările pe modelul matematic dezvăluie un număr mare de desfăşurări posibile ale procesului social analizat, în variantele cele mai probabile. Aceasta apare în urma experimentelor efectuate pe acest model şi relevă căile posibile de intervenţie în procesul obiectiv, conform legilor sale de dezvoltare.

Utilitatea matematicii pentru cercetările sociologice este însă insuficient lămurită, fiind uneori chiar contestată. Controversele numeroase apărute se datorează atît acelora care, cultivîndu-şi rutina, se opun *matematizării*, cît şi militanţilor utilizării matematicii în domeniul sociologiei, care nu au învins încă greutăţile în analiza informaţiilor sociologice cu metode matematice.

Problema principală care stă în faţa sociologiei matematice este tocmai formularea matematică a unei game largi de fenomene şi procese sociale, care să facă posibilă rezolvarea unei probleme complicate cu ajutorul calculatoarelor electronice.

Scanned with OKEN Scanner

Uneori este emisă părerea că matematica poate servi numai la studierea aspectului cantitativ al proceselor sociale şi, ca atare, aplicarea sa este îngrădită de însăşi limitele posibilităţilor de cuantificare a unor aspecte ale realităţii sociale. Această părere este deja perimată, deoarece în ultima jumătate de veac categoria cantităţii a fost înlocuită prin categoria structurii. Matematica modernă nu mai este subordonată aspectelor cantitative, deoarece şi-a dezvoltat baza teoretică şi în domeniul variabilelor discontinue, care sugerează apropieri surprinzătoare de problemele studiate de ştiinţele umane. Mai mult, discontinuitatea este tocmai una din caracteristicile esenţiale ale raportului între ansambluri calitative cum sînt cele studiate de sociologie. Această idee a deschis un cîmp larg de aplicare a *matematicii calitative* în studiul fenomenelor şi proceselor sociale, matematică cu perspective excepţionale în domeniul sociologiei.

Matematica utilizată drept instrument al sociologiei poate constitui o reprezentare atît a aspectului cantitativ, cît şi a celui calitativ al fenomenului sau procesului cuprins în sfera obiectului acestei ştiinţe.

Formularea matematică a proceselor sociale, implicînd folosirea limbajului matematic în ştiinţele sociale, se impune ca un *mijloc de cunoaştere*. Traducerea proceselor sociale în limbajul matematicii, prin intermediul unor modele matematice care le corespund, ne oferă posibilitatea să dezvăluim pe cale deductivă multe din laturile şi manifestările lor care au fost doar vag intuite.

Teoretic privind, orice raţionament corect în domeniul sociologiei se poate formaliza matematic; orice proces social poate fi tradus în termenii unui model matematic sau logic-matematic, dacă admite o reprezentare calitativă adecuată, formulabilă într-un limbaj simbolic, ca de pildă cel matematic.

În general, tratarea matematică a unei probleme sociologice nu numai că nu o complică, ci o simplifică. În dome-

Scanned with OKEN Scanner

niul sociologiei există numeroase probleme care numai matematic pot fi formulate complet şi concludent, ca de pildă structura unui proces social. În practică, garanţia corectitudinii logice a reprezentării unei astfel de structuri o putem obţine numai prin construirea unui model simbolic, de preferinţă matematic, apt să simuleze transformările posibile în structura socială luată în studiu.

Apariţia conceptelor de *model* şi *structură* este strîns legată de dezvoltarea şi perfecţionarea metodelor matematice aplicate în ştiinţele sociale. De pildă, o extindere remarcabilă au căpătat aplicaţiile teoriei mulţimilor, teoriei grafelor şi a algebrei matriciale în studiul structurilor sociale foarte diferite : structura grupului uman, structura procesului social, structura relaţiilor de rudenie ale unui grup de indivizi, structura fluxului de mobilitate a populaţiei etc. Reprezentarea unor astfel de structuri prin grafe şi matrici asociate lor contribuie la efectuarea unei analize profunde a structurii sociale considerate, deoarece numai prin descoperirea înlănţuirii cauzale a caracteristicilor esenţiale putem descifra adînc o astfel de structură.

Unele metode şi modele matematice utilizate în ştiinţele umane, elaborate prin intermediul unor teorii matematice ca : teoria mulţimilor, teoria grafelor, teoria proceselor stocastice, teoria informaţiei, etc., pot lărgi sfera înţelegerii şi cunoaşterii ştiinţifice a unor procese sociale deosebit de complexe cum ar fi : procesul de migraţie a populaţiei de la sat la oraş, procesul de migraţie interzonală a populaţiei de pe teritoriul ţării, procesul de transmitere a culturii din mediul urban în cel rural, procesul de mobilitate socială şi profesională, procesul de stratificare, procesul de urbanizare etc.

Procesele sociale pe care le studiază sociologia ca ştiinţă sînt determinate de o serie de factori economici, biologici geografici, psihici, juridici etc., care pot fi puşi în evidenţă cu ajutorul unor *indicatori definitorii* (descriptivi), *indicatori corelaţionali* şi *indicatori inferenţiali*. Ansamblul

Scanned with OKEN Scanner

acestor indicatori poate fi asociat unui proces social oarecare; relaţiile dintre aceşti indicatori caracterizează structura acestui proces.

Relaţiile statistice dintre caracteristicile unei structuri sociale pot fi puse în evidenţă cu ajutorul indicatorilor corelaţionali şi inferenţiali. De aceea, analiza corelaţiilor statistice şi a inferenţelor cauzale dintre variabile prezintă o valoare excepţională pentru studiul proceselor sociale.

O contribuţie importantă o poate aduce sociologiei formularea unor *modele de corelaţii şi de relaţii cauzale* ale proceselor sociale, adică construirea unor asemenea reprezentări care să permită analiza şi interpretarea pe baze ştiinţifice a unor structuri sociale deosebit de complexe. Utilizarea unor astfel de reprezentări ar putea deschide un cîmp larg de aplicare a statisticii şi matematicii în cercetările sociologice.

Pentru studiul proceselor sociale cu ajutorul modelelor şi metodelor matematice, relaţiile cauzale au o valoare mult mai mare decît corelaţiile statistice, deoarece primele stabilesc o dependenţă pe cînd ultimele o asociere între variabile. De aceea, punerea la punct a aparatului matematic necesar elaborării *modelelor cauzale* ale proceselor sociale, constituie o problemă centrală a sociologiei matematice. De rezolvarea acestei probleme ar depinde în viitor succesele modelării matematice a proceselor sociale.

O condiţie esenţială pentru succesul modelării matematice a proceselor sociale este aceea că întotdeauna trebuie efectuată *o analiză multivariată a structurilor cauzale.* Această analiză este superioară în mod evident, deoarece exclude în mare măsură, posibilitatea ca dependenţele dintre variabile să fie „alterate" de alte influenţe exterioare.

Elaborarea unor modele cauzale ale proceselor sociale contribuie la :

Scanned with OKEN Scanner

a) sporirea eficacității metodelor de analiză matematică aplicate în cercetările sociologice;

b) optimizarea unor fenomene sau procese sociale;

c) aplicarea calculatoarelor electronice în cercetările sociologice.

Elaborarea unor modele ale unor structuri sociale concrete permite *dirijarea* pe baze științifice a unor procese sociale deosebit de complexe, cum ar fi: procesul de migrație între diviziunile teritorial administrative (județe) cu nivel diferit de dezvoltare industrială și urbană. Astfel prin direcționarea unor factori sociali-economici poate fi dirijat, controlat și influențat acest proces deosebit de actual. În acest scop, modelele cauzale pot oferi baza metodologică previzională pentru elaborarea planurilor de perspectivă în cadrul unităților centrale de planificare și control.

Modelarea matematică și tehnica electronică de calcul pot aduce în viitor o contribuție importantă la *perfecționarea activităților umane*. De pildă, aplicațiile matematicii pot fi extinse cu succes în studiul structurii și dinamicii grupurilor umane și în știința despre comportamentul uman. Activitatea grupurilor umane poate fi *standardizată* prin introducerea unor norme raționale de comportament uman în cadrul acestora.

Aplicațiile algebrei moderne în științele umane pot fi numeroase: matricele sociometrice și rețelele de comunicație utilizate în studiul grupurilor mici; aplicațiile proceselor stocastice în genetică; aplicațiile lanțurilor Markov în sociologie, psihologie și biologie; modele de comportament uman; modele de creștere economică etc.

Unele teorii matematice s-au născut mai întîi în contact cu probleme ale științelor umane. Astfel s-au pus bazele teoriei jocurilor și teoriei matematice a gramaticii. Dintre acestea, teoria jocurilor necompetitive are cele mai mari șanse de a exprima veridic realitatea socială.

Perfecționarea și extinderea aplicațiilor matematicii în domeniul sociologiei este corelată foarte strîns cu folo-

Scanned with OKEN Scanner

sirea calculatoarelor electronice în acest domeniu. Aplicaţiile tehnicii electronice de calcul aduc deja foloase remarcabile cercetărilor sociologice prin : extinderea clasificării automate a informaţiilor sociologice, analiza structurii latente, simularea proceselor sociale etc. De exemplu, în psihologie îndeosebi există cîteva aplicaţii interesante de simulare pe calculator a activităţii de inducţie, impunînd maşinii, diverse „cunoştinţe" şi conduite posibile.

Se poate afirma că metodele de simulare a proceselor sociale deosebit de complexe pot aduce o contribuţie însemnată în domeniul sociologiei, îndeosebi acolo unde experimentarea este imposibilă şi unde, complexitatea ipotezelor face traducerea matematică foarte dificilă. Fără îndoială că metodele şi modelele matematice care extind sfera analizei acţiunilor şi deciziilor umane capătă o dezvoltare tot mai mare în domeniul sociologiei. Această dezvoltare care rezultă din nevoile crescînde ale sociologiei, constituie un cîştig pentru dezvoltarea ei ca ştiinţă.

Traducerea proceselor sociale în cifre şi simboluri permite studiul unui număr foarte mare de informaţii sociologice, confruntarea caracteristicilor respective cu mare precizie şi aprofundarea analizei matematice. Dacă sociologii vor introduce maximum de cuantificare în cercetările lor va creşte rolul aplicaţiilor matematicii şi calculatoarelor electronice în domeniul sociologiei.

În general, între rezultatele ce se pot obţine cu ajutorul metodelor clasice de cercetare şi a tehnicilor matematice moderne, există o diferenţă foarte mare. Cu ajutorul matematicii se poate adînci analiza proceselor sociale dincolo de limitele observaţiei nemijlocite a cercetătorului. Cu toate acestea, este posibil ca unele domenii deosebit de interesante ale sociologiei să nu se preteze mult timp la cuantificarea şi la analiza matematică. Ritmul dezvoltării şi perfecţionării metodelor matematice aplicate în cercetările sociologice este foarte strîns legat de succesele viitoare ale matematicii şi ale sociologiei. Din întrepătrunderea acestor două ştiinţe se va putea dezvolta *sociologia matematică*, a cărei bază a fost pusă deja în unele ţări

Scanned with OKEN Scanner

(S. U. A., Anglia, Franţa etc.). De altfel, această disciplină relativ nouă a început să fie predată şi într-o serie de universităţi din aceste ţări.

Cu toate că sociologia este susceptibilă de matematizare, domeniul de aplicabilitate a metodelor matematice a fost multă vreme destul de restrîns datorită pe de o parte, complexităţii problemelor specifice acestei ştiinţe, iar pe de altă parte, insuficienţei unor teorii matematice, care să fie puse în serviciul nevoilor practice ale cercetării sociologice. Cu toate acestea, în ultimii 10—15 ani s-a înregistrat însă un ritm extrem de accelerat al aplicaţiilor matematicii în domeniul sociologiei, graţie elaborării unor teorii şi metode matematice, cum ar fi : teoria mulţimilor, teoria grafelor, teoria jocurilor, teoria informaţiei etc.

Atîta timp cît sociologia va face apel la matematică pe o scară mai largă, ea va da un impuls considerabil perfecţionării sau chiar creării unor teorii matematice noi. La rîndul lor, teoriile matematice noi vor spori amploarea şi eficienţa metodelor de cercetare aplicate în domeniul sociologiei.

În stadiul actual este de neconceput dezvoltarea teoretică şi metodologică a sociologiei sau chiar rezolvarea unei probleme complexe în acest domeniu fără o bază matematică bine pusă la punct şi de nivel ridicat.

Mulţi sociologi sînt de părere că pentru necesităţile sociologiei este suficient să se folosească metode statistico-matematice clasice, pe alocuri insuficiente, introduse ca auxiliar al unor metode şi tehnici de cercetare sociologică. Această rezolvare este desigur mai uşoară şi mai accesibilă, dar are dezavantajul că în condiţiile complexităţii fenomenelor şi proceselor sociale, tratarea statistico-matematică cu ajutorul metodelor clasice poate conduce, în anumite situaţii, la rezultate eronate, ca de pildă calculul corelaţiei şi regresiei simple în cazul unei structuri complexe. Desigur că, în cercetările sociologice, trebuie să elaborăm modele matematice care să reflecte procesele sociale analizate în toată complexitatea lor, evitîndu-se pe cît posibil studiul unor „fracţiuni“ de structuri sociale

concrete. De aceea, în studiul proceselor sociale complexe, utilizarea matematicii moderne şi a calculatoarelor electronice este cert superioară şi indispensabilă.

Utilizarea matematicii în studiul proceselor sociale întîmpină o serie de dificultăţi însemnate. Aceste dificultăţi sînt puse în strînsă legătură cu faptul că în lumea reală, procesele sociale nu sînt izolate, ci fac parte dintr-un sistem complex. Datorită acestui fapt, este dificil să se izoleze şi să se descopere procese sociale separate. Greutăţile legate de evidenţierea unor procese sociale concrete face extrem de dificilă sintetizarea acestora într-un model matematic. Problema desprinderii proceselor sociale izolate din sistemul complex al structurii întregii societăţi trebuie să fie rezolvată de către sociologi; numai în asemenea condiţii va fi posibilă realizarea unui ansamblu de modele matematice ce pot fi rezolvate cu ajutorul calculatoarelor electronice, pentru a le descoperi comportamentul în condiţii diferite.

Este destul de dificil de a prezenta o imagine clară asupra tendinţelor viitoare ale extinderii aplicaţiilor matematicii şi calculatoarelor electronice în cercetările sociologice. Ceea ce se poate afirma în prezent este că în cercetările sociologice contemporane există anumite nevoi: *ameliorarea metodelor de analiză şi a mijloacelor previzionale în cercetările sociologice.*

Factorii care vor influenţa, probabil, dezvoltarea metodelor matematice în cercetările sociologice vor fi:

— dezvoltarea şi perfecţionarea calculatoarelor electronice;
— dezvoltarea tehnicilor matematice;
— dezvoltarea cercetărilor sociologice de mare complexitate;
— ameliorarea mijloacelor previzionale şi de planificare socială;
— dezvoltarea procedeelor de cercetare şi de culegere a informaţiilor (extinderea sondajelor contextuale şi a sondajelor panel);

Scanned with OKEN Scanner

— dezvoltarea învăţămîntului de metodologie, de sociologie, de statistică şi de matematică;
— interesul crescînd al matematicienilor în aplicarea matematicii la ştiinţele umane.

Fără îndoială că succesele viitoare ale aplicaţiilor matematicii şi calculatoarelor electronice în cercetările sociologice vor contribui la dezvoltarea unei metodologii eficace de cercetare a proceselor sociale. Desigur, viitorul s-ar putea să ne ofere perspective excepţionale, dacă avem în vedere faptul că problemele cele mai abstracte ale teoriei sociologice, relevă ele însele o metodologie matematică exactă.

Probabil că în anii următori vom asista la o dezvoltare importantă a metodologiei matematice în cercetările sociologice, ca urmare a faptului că dezvoltarea sociologiei ca ştiinţă va cere matematici din ce în ce mai complexe, dată fiind complexitatea fenomenelor şi proceselor sociale. De pildă, dezvoltarea teoriei generale a sistemelor, a teoriei structurilor etc. vor contribui la extinderea aplicaţiilor matematicii şi a calculatoarelor electronice în domeniul sociologiei. Fără îndoială că dezvoltarea unor instrumente ca statistica matematică, teoria jocurilor, teoria informaţiei, sau a ciberneticii, va constitui un fond extrem de valoros pentru punerea la punct a metodologiei matematice de cercetare sociologică.

Lucrarea *Modelarea matematică a proceselor sociale* relevă faptul că între matematică şi sociologie se poate stabili o legătură tot mai strînsă, care să contribuie la lărgirea necontenită a aplicaţiilor matematicii moderne şi a calculatoarelor electronice în cercetările sociologice.

Metodele cantitative utilizate în cercetările sociologice se impun şi se perfecţionează tot mai mult în vastul cîmp al problemelor legate de modelarea proceselor sociale, ca unul dintre cele mai importante mijloace de cunoaştere ştiinţifică în acest domeniu. Dacă ne gîndim că în domeniul sociologiei raţionăm logic, este cel mai util de a da o formă matematică raţionamentului nostru, chiar dacă rezolvăm problema aproximativ. De altfel, dacă mate-

Scanned with OKEN Scanner

matizăm o problemă nu urmărim altceva decît de a da un model aproximativ al procesului social studiat. După aceea, modelul poate să nu fie numeric, ci grafic (prezentat sub forma unui graf).

Desigur că, în viitor, exactitatea metodelor matematice va spori, ca urmare a formulării unui complex de modele eficace ale proceselor sociale, a creării unui sistem unic de informaţii sociale şi a unei reţele unice de stat de centre de calcul. Numai în asemenea condiţii va creşte foarte mult importanţa metodologiei matematice pentru cercetările sociologice concrete, a sociologiei matematice ca ştiinţă cu perspective rodnice.

Scanned with OKEN Scanner

CUPRINS

	Pag.
1. Preocupări actuale de modelare matematică a proceselor sociale	9
1.1. Introducere	9
1.2. Aportul statisticii matematice în cercetarea proceselor sociale	11
1.3. Formalizarea informațiilor sociologice	12
1.4. Dezvoltarea matematică a modelelor	16
1.5. Confruntarea predicțiilor modelului matematic cu informațiile statistice	28
1.6. Concluziile cercetărilor de sociologie matematică	30
1.7. Aplicațiile tradiționale ale matematicii în științele sociale	35
1.8. Diferențierea funcțiilor matematicii	38
1.9. Perspectivele aplicațiilor matematicii în domeniul cercetărilor sociologice	40
1.10. Concluzii	45
Bibliografie	46
2. Modele inductive	48
2.1. Introducere	48
2.2. Conceptele de proces, structură și model	50
2.3. Unele aplicații generale ale teoriei mulțimilor și ale teoriei grafelor în analiza proceselor sociale	56
2.4. Reprezentarea relațiilor de dependență cu ajutorul grafelor orientate	58
2.5. Elemente de bază ale grafului orientat	64
2.6. Graful și drumurile într-un graf orientat	65
2.7. Matricea asociată unui graf	71
2.8. Metoda de cercetare a depărtărilor într-un graf	77

Scanned with OKEN Scanner

Pag.

2.9. Un model inductiv aplicat în studiul procesului de migrație a populației 82
2.10. Un alt exemplu : studiul procesului de transmitere a culturii din mediul urban în cel rural 91
2.11. Analiza relațiilor dintre membrii unui grup social 102
2.12. Concluzii 106

Bibliografie 109

3. Modele de analiză corelațională 111

3.1. Introducere 111
3.2. Natura relațiilor statistice 114
3.3. Clasificarea relațiilor statistice 117
3.4. Principii teoretice ale analizei relațiilor statistice 118
3.5. Metode elementare de calcul a corelațiilor simple . . . 119
3.6. Metode analitice de calcul a corelațiilor simple 129
3.7. Corelația rangurilor 135
3.8. Asocierea caracteristicilor calitative 140
3.9. Corelația parțială 145
3.10. Corelația multiplă 147
3.11. Determinația multiplă 149
3.12. Verificarea semnificației coeficientului de corelație multiplă 151
3.13. Concluzii 152

Bibliografie 155

4. Modele de analiză cauzală 159

4.1. Introducere 159
4.2. Noțiunile de corelație statistică și de relație cauzală . . . 162
4.3. Deducerea relațiilor cauzale din relațiile statistice 166
4.4. Analiza dependenței între două variabile 175
4.5. Coeficientul de dependență al lui Wright 184
4.6. Modele liniare de analiză a structurilor cauzale complexe 188
4.7. Calculul regresiei liniare multiple 193
4.8. Un exemplu de analiză a relațiilor de dependență între nivelul de instruire, gradul de calificare și productivitatea muncii 200

Scanned with OKEN Scanner

Pag.

4.9. Un alt exemplu : modelul cauzal al procesului de urbanizare 204
4.10. Concluzii 211
Bibliografie 212

5. Modele de analiză factorială şi multifactorială 215

5.1. Introducere 215
5.2. Metoda de analiză factorială a lui Spearman 216
5.3. Un exemplu de analiză a factorului de progres tehnic înregistrat în economia naţională 220
5.4. Un exemplu de construire a unui indicator sintetic al nivelului de trai al naţiunilor 223
5.5. Un exemplu de analiză factorială a relaţiei dintre caracteristicile social-politice şi produsul global pe locuitor 227
5.6. Un exemplu de analiză factorială a poziţiei social-economice a agricultorilor 233
5.7. Analiza structurii latente 239
5.8. Concluzii 244
Bibliografie 245

Încheiere. 247

Scanned with OKEN Scanner

MTH 2100

Scanned with OKEN Scanner

ERATĂ

La pag.:	*rîndul:*	*în loc de:*	*se va citi:*
52	9 de jos	metodelor	modelelor
57	11 de sus	Y	y
64	4 de sus	$x_1,\ x_3$	x_i, x_j
186	1 de sus	$\frac{\sigma_1^2}{\sigma_2^2}$	$\frac{\sigma_1}{\sigma_2}$
204	9 de sus	$b_{12}x_1 + \ldots + e_i = e_i,$ și nu sub forma $x_1 = a_{12}x + \ldots + e_i$	$b_{12}x_1 + x_2 = e_2,$ și nu sub forma $x_2 = a_{12}\ x_1 + e_2$

Redactor: BARBU BERCEANU
Tehnoredactor: MARILENA DAMASCHINOPOL

Bun de tipar 24. 04. 972. Hîrtie tipar înalt ilustr., format 16/58 × 84 de 80 g/m². Coli de tipar 16,75.

C. Z. pentru bibliotecile mari { 301:5192 / 519.2:301

C. Z. pentru biblioteci mici 301:519

Întreprinderea poligrafică Informaţia, Bucureşti, str. Brezoianu nr. 23 — 25
Republica Socialistă România

Scanned with OKEN Scanner

Lei 7,75

În BIBLIOTHECA SOCIOLOGICA

a apărut, în seria Monografii :

1. **Procesul de urbanizare în R. S. România, Zona Slatina-Olt,** colectivul de autori : Aculin CAZACU, Stela CERNEA, Septimiu CHELCEA, Miron CONSTANTINESCU, Virgil CONSTANTINESCU, Gheorghe CONSTANTIN, Ștefan COSTEA, Natalia DAMIAN, Ion DRĂGAN, Florian GEORGESCU, Pompiliu GRIGORESCU, Henri H. STAHL, Petre STĂNCULESCU și Leon ȚOPA, sub redacția prof. univ. Miron CONSTANTINESCU și prof. univ. H. H. STAHL, 1970, 398 (—399) p.

Vor apărea :

Maria FULEA și Maria COBIANU, **Organizarea muncii în cooperativele agricole de producție**

M. MERFEA, **Procesul de urbanizare în R. S. România (Zona Vaslui)**

M. ZACHIU, **Utilizarea forței de muncă în județul Timiș**

Constantin IONESCU, **Aspecte ale bugetului de timp al oamenilor muncii**

I. MATEI, **Dezvoltarea corelată a mediului urban și a celui rural**

Oscar HOFFMAN, **Sociologia valorii**

Aneta RADU-ALMĂȘAN, **Lichidarea opoziției și apropierea treptată între oraș și sat în etapa actuală de dezvoltare a țării noastre**

Scanned with OKEN Scanner

Lei 7,75

În BIBLIOTHECA SOCIOLOGICA

a apărut, în seria Monografii :

1. **Procesul de urbanizare în R. S. România, Zona Slatina-Olt,** colectivul de autori : Aculin CAZACU, Stela CERNEA, Septimiu CHELCEA, Miron CONSTANTINESCU, Virgil CONSTANTINESCU, Gheorghe CONSTANTIN, Ștefan COSTEA, Natalia DAMIAN, Ion DRĂGAN, Florian GEORGESCU, Pompiliu GRIGORESCU, Henri H. STAHL, Petre STĂNCULESCU și Leon TOPA, sub redacția prof. univ. Miron CONSTANTINESCU și prof. univ. H. H. STAHL, 1970, 398 (—399) p.

Vor apărea :

Maria FULEA și Maria COBIANU, **Organizarea muncii în cooperativele agricole de producție**

M. MERFEA, **Procesul de urbanizare în R. S. România (Zona Vaslui)**

M. ZACHIU, **Utilizarea forței de muncă în județul Timiș**

Constantin IONESCU, **Aspecte ale bugetului de timp al oamenilor muncii**

I. MATEI, **Dezvoltarea corelată a mediului urban și a celui rural**

Oscar HOFFMAN, **Sociologia valorii**

Aneta RADU-ALMĂȘAN, **Lichidarea opoziției și apropierea treptată între oraș și sat în etapa actuală de dezvoltare a țării noastre**

Scanned with OKEN Scanner